

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

## OM protein - protein search, using sw model

Run on: May 23, 2005, 18:56:58 ; Search time 167 Seconds  
(without alignments)  
1185.756 Million cell updates/sec

Title: us-10-081-816-12

Perfect score: 2645 1 MRPSGRKVKHGQNSGHS.....TTTVGTTLPPIMKLGRKG 512

Sequence: 1 MRPSGRKVKHGQNSGHS.....TTTVGTTLPPIMKLGRKG 512

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneeqp19808:\*
- 2: geneeqp19908:\*
- 3: geneeqp20008:\*
- 4: geneeqp20018:\*
- 5: geneeqp20028:\*
- 6: geneeqp20038:\*
- 7: geneeqp20048:\*
- 8: geneeqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2645	100.0	512	4	ABB70017 Drosophila
2	2645	100.0	512	5	AAB30492 Fruit fly
3	405	15.3	454	4	ABB68814 Drosophila
4	390	14.7	364	4	AAB75197 Drosophila
5	390	14.7	364	5	AAB30504 Fruit fly
6	169	6.4	33	5	AAB30548 Fruit fly
7	160	6.0	433	4	AAB75232 Drosophila
8	135.5	5.1	369	5	AAB30509 Fruit fly
9	132.5	5.0	736	4	ABB67525 Drosophila
10	132.5	5.0	736	5	AAB30497 Fruit fly
11	132.5	5.0	1299	4	ABB70025 Drosophila
12	130.5	4.9	372	4	AAB75208 Drosophila
13	130.5	4.9	372	5	AAB30510 Fruit fly
14	129	4.9	436	4	ABB68743 Drosophila
15	129	4.9	436	5	AAB30491 Fruit fly
16	127	4.8	348	5	AAB75207 Drosophila
17	122	4.6	410	4	ABB71166 Drosophila
18	122	4.6	410	5	AAB30481 Fruit fly
19	122	4.6	414	4	AAB75193 Drosophila
20	120	4.5	450	5	AAB30537 Fruit fly
21	119	4.5	477	4	AAB75227 Drosophila
22	118.5	4.5	2008	4	AAB75227 Drosophila
23	118.5	4.5	2008	5	AAB75227 Drosophila
24	118.5	4.5	2008	6	ABR43636 Human
25	118.5	4.5	2008	7	ADC79417 Human

26	115	4.3	367	5	AAB30493	Aae30493 Fruit fly
27	115	4.3	817	4	ABB70027	Abb70027 Drosophila
28	114	4.3	693	4	ABB47271	Abb47271 Enterococ
29	113	4.3	250	4	ABB67792	Abb67792 Drosophila
30	113	4.3	378	7	ADD15295	Add15295 Fruit fly
31	112.5	4.3	365	5	AAB30536	Aae30536 Fruit fly
32	112.5	4.3	408	4	AAB75217	Abb75217 Drosophila
33	112.5	4.3	698	6	ABM67593	Abm67593 Phototrab
34	112.5	4.3	1227	4	ABB68608	Abb68608 Drosophila
35	110.5	4.2	408	5	AAB30517	Aae30517 Fruit fly
36	110.5	4.2	2073	6	ABR43629	Abt43629 Mouse CLA
37	110	4.2	498	4	ABB63071	Abb63071 Drosophila
38	110	4.2	498	5	AAB30495	Aae30495 Fruit fly
39	109.5	4.1	327	8	ADK16371	Adk16371 Nanoarcha
40	108	4.1	381	4	AAB75209	Abb75209 Drosophila
41	108	4.1	381	5	AAB30511	Aae30511 Fruit fly
42	107.5	4.1	408	5	ABP26074	Abp26074 Streptoco
43	107.5	4.1	346	6	ABU49652	Abu49652 Protein e
44	107	4.0	404	5	AAB30532	Aae30532 Fruit fly
45	106.5	4.0	390	4	AAB75221	Abb75221 Drosophila

## ALIGNMENTS

RESULT 1  
ABB70017 standard; protein; 512 AA.  
ID ABB70017

AC ABB70017;

DE 26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 36843.

Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US009231.

23-MAR-2000; 2000US-0191637P.

11-JUL-2000; 2000US-00614150.

(PERKE ) PE CORP NY.

Venter JC, Adams M, Li FWD, Myers EW;

WPI; 2001-656860/75.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Disclosure; SEQ ID NO 36843; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AAB57737-AB122072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 512 AA;  
 Query Match 100.0%; Score 2645; DB 4; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-302;  
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPSGEKVYKGGHSGHSGHSGMANYRRKKGDAVFLNAKPLNSANAQAYLYGVKYSI 60  
 DB 1 MRPSGEKVYKGGHSGHSGHSGMANYRRKKGDAVFLNAKPLNSANAQAYLYGVKYSI 60  
 QY 61 GLAERLDADYEAPPLDRKSSDSTASNNPEKPSVYRNIDINMFLRIIGVLPYRHGP 120  
 DB 61 GLAERLDADYEAPPLDRKSSDSTASNNPEKPSVYRNIDINMFLRIIGVLPYRHGP 120  
 QY 121 ARAKPEMNSASFTYSVVFVLLACYGYVANNRHHVRSLSGPFBEAVIAYFLVNIILPI 180  
 DB 121 ARAKPEMNSASFTYSVVFVLLACYGYVANNRHHVRSLSGPFBEAVIAYFLVNIILPI 180  
 QY 181 MIIPILWEARKIAKLFNDMDFEVLYYQISGHSLEPLKORAVYIAYVLPILSVLSVI 240  
 DB 181 MIIPILWEARKIAKLFNDMDFEVLYYQISGHSLEPLKORAVYIAYVLPILSVLSVI 240  
 QY 241 THVTMSDLINOVVPCILDNTAMLGAMWFLICEAMSTIAHLAERFQKALKHIGPAM 300  
 DB 241 THVTMSDLINOVVPCILDNTAMLGAMWFLICEAMSTIAHLAERFQKALKHIGPAM 300  
 QY 301 VADYRVLMLRLSKLTRDTGNALCTFVFMSTLYLFFIITLSYGLMSQSEGGIIGLT 360  
 DB 301 VADYRVLMLRLSKLTRDTGNALCTFVFMSTLYLFFIITLSYGLMSQSEGGIIGLT 360  
 QY 361 ITALMNIGLFLYICDEAHYASVNVRTNPOKLLMELMNMNSDAQEIMFLRATEMNS 420  
 DB 361 ITALMNIGLFLYICDEAHYASVNVRTNPOKLLMELMNMNSDAQEIMFLRATEMNS 420  
 QY 421 TINCQGFDPVNRKTLFRGLITTWVTVYLVLLQFOISIPDTKDGSEGANNTTVDPVWDSLD 480  
 DB 421 TINCQGFDPVNRKTLFRGLITTWVTVYLVLLQFOISIPDTKDGSEGANNTTVDPVWDSLD 480  
 QY 481 NDMSLMGASTLSTTVGTTLPPPIKLGKRGK 512  
 DB 481 NDMSLMGASTLSTTVGTTLPPPIKLGKRGK 512

RESULT 2  
 AAE30492  
 ID AAE30492 standard; protein; 512 AA.  
 XX  
 AC AAE30492;  
 XX  
 DT 24-FEB-2003 (first entry)  
 XX  
 DE Fruit fly gustatory receptor protein, Gr63f1.  
 XX  
 KM Fruit fly, gustatory receptor; Gr; odorant receptor; Or; pesticide;  
 XX Gr63f1.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN MO200268593-AA2.  
 PD 06-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-US005414.  
 XX  
 PR 23-FEB-2001; 2001US-0271319P.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Axel R, Scott K;  
 XX  
 DR MPI; 2002-698668/75.  
 XX  
 PT Novel nucleic acid encoding insect gustatory or odorant receptor protein

PT useful for identifying a compound which specifically binds to the  
 receptor for controlling a pest population in an area.  
 PS Claim 3; Page 160-162; 264pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid encoding an insect  
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises  
 CC seven transmembrane domains and a C-terminal domain comprising  
 CC consecutive amino acids. The invention is useful for identifying a  
 CC compound which activates the insect receptor or inhibits the activity of  
 CC the insect receptor. The purified insect receptor protein is embedded in  
 CC a lipid bilayer. The invention is sprayed for combating ingestion of  
 CC crops by pest insects, combating disease-carrying insects in an area and  
 CC controlling a pest population in an area. The invention is useful for  
 CC detecting the presence of insect gustatory or odorant receptor and for  
 CC inhibiting the function of the receptor in humans or animals or in  
 CC biological fluids isolated from them. The invention is also useful for  
 CC identifying or isolating other insect receptors and for combating pest  
 CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)  
 CC protein  
 XX

SQ Sequence 512 AA;  
 Query Match 100.0%; Score 2645; DB 5; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-302;  
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPSGEKVYKGGHSGHSGHSGMANYRRKKGDAVFLNAKPLNSANAQAYLYGVKYSI 60  
 DB 1 MRPSGEKVYKGGHSGHSGHSGMANYRRKKGDAVFLNAKPLNSANAQAYLYGVKYSI 60  
 QY 61 GLAERLDADYEAPPLDRKSSDSTASNNPEKPSVYRNIDINMFLRIIGVLPYRHGP 120  
 DB 61 GLAERLDADYEAPPLDRKSSDSTASNNPEKPSVYRNIDINMFLRIIGVLPYRHGP 120  
 QY 121 ARAKPEMNSASFTYSVVFVLLACYGYVANNRHHVRSLSGPFBEAVIAYFLVNIILPI 180  
 DB 121 ARAKPEMNSASFTYSVVFVLLACYGYVANNRHHVRSLSGPFBEAVIAYFLVNIILPI 180  
 QY 181 MIIPILWEARKIAKLFNDMDFEVLYYQISGHSLEPLKORAVYIAYVLPILSVLSVI 240  
 DB 181 MIIPILWEARKIAKLFNDMDFEVLYYQISGHSLEPLKORAVYIAYVLPILSVLSVI 240  
 QY 241 THVTMSDLINOVVPCILDNTAMLGAMWFLICEAMSTIAHLAERFQKALKHIGPAM 300  
 DB 241 THVTMSDLINOVVPCILDNTAMLGAMWFLICEAMSTIAHLAERFQKALKHIGPAM 300  
 QY 301 VADYRVLMLRLSKLTRDTGNALCTFVFMSTLYLFFIITLSYGLMSQSEGGIIGLT 360  
 DB 301 VADYRVLMLRLSKLTRDTGNALCTFVFMSTLYLFFIITLSYGLMSQSEGGIIGLT 360  
 QY 361 ITALMNIGLFLYICDEAHYASVNVRTNPOKLLMELMNMNSDAQEIMFLRATEMNS 420  
 DB 361 ITALMNIGLFLYICDEAHYASVNVRTNPOKLLMELMNMNSDAQEIMFLRATEMNS 420  
 QY 421 TINCQGFDPVNRKTLFRGLITTWVTVYLVLLQFOISIPDTKDGSEGANNTTVDPVWDSLD 480  
 DB 421 TINCQGFDPVNRKTLFRGLITTWVTVYLVLLQFOISIPDTKDGSEGANNTTVDPVWDSLD 480  
 QY 481 NDMSLMGASTLSTTVGTTLPPPIKLGKRGK 512  
 DB 481 NDMSLMGASTLSTTVGTTLPPPIKLGKRGK 512

RESULT 3  
 ABB68814  
 ID ABB68814 standard; protein; 454 AA.  
 XX  
 AC ABB68814;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 33234.

[illegible][illegible]

```

Db      165 CSIMYINCNAFGTASRALSDALQTTIRGEKPAQKLTREYRHLWVDLSHMMQOUGRAYSNNY 224
QY      327 VEMSLYLFPIITLSIYGLMSQ--LSBEGFGIKDIGITLALMNIIGLFYICDEAHYASVNR 385
Db      225 GMVCLVIFFTTITATYGSIEIIDHGATYKEVGLFVIVCYCGLLYIICNEAHYASRKVG 284
QY      386 TNFOKKLWVEIMNMNSDAQTEINMPLRATENMPTINGCFEPDVNRTLFGKILTTMTY 445
Db      285 LDFQTKLAININTAVDAATQKEVEMLVAINKNPPIIMNDGVANINREILTTNISMTATY 344
QY      446 LVVLLQFOIS 455
Db      345 LVVLLQFKIT 354

```

## RESULT 5

AAE30504 standard; protein; 364 AA.

```

AC      AAE30504;
XX      24-FEB-2003 (first entry)
DT      24-FEB-2003 (first entry)
XX      Fruit fly gustatory receptor protein, Gr21D1.
XX      Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;
XX      Gr21D1.
XX      Drosophila melanogaster.
XX      WO200268593-A2.
XX      06-SEP-2002.
XX      22-FEB-2002; 2002WO-US005414.
XX      23-FEB-2001; 2001US-0271319P.
XX      (UYCO ) UNITV COLUMBIA NEW YORK.
XX      Axel R, Scott K;
XX      WPI; 2002-698668/75.
XX      Novel nucleic acid encoding insect gustatory or odorant receptor protein
XX      useful for identifying a compound which specifically binds to the
XX      receptor for controlling a pest population in an area.
XX      Disclosure; Page 183-185; 264pp; English.
XX      The invention relates to an isolated nucleic acid encoding an insect
XX      gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
XX      seven transmembrane domains and a C-terminal domain comprising
XX      consecutive amino acids. The invention is useful for identifying a
XX      compound which activates the insect receptor or inhibits the activity of
XX      the insect receptor. The purified insect receptor protein is embedded in
XX      a lipid bilayer. The invention is sprayed for combating ingestion of
XX      crops by pest insects, combating disease-carrying insects in an area and
XX      controlling a pest population in an area. The invention is useful for
XX      detecting the presence of insect gustatory or odorant receptor and for
XX      inhibiting the function of the receptor in humans or animals or in
XX      biological fluids isolated from them. The invention is also useful for
XX      identifying or isolating other insect receptors and for combating pest
XX      nuisances. The present sequence is fruit fly gustatory receptor (Gr)
XX      protein
XX      Sequence 364 AA;

```

Query Match 14.7%; Score 390; DB 5; Length 364;  
 Best Local Similarity 27.3%; Pred. No. 3.6e-36;  
 Matches 101; Conservative 77; Mismatches 152; Indels 40; Gaps 9;

```

QY      110 IGVLPVIRHGP---ARAFENNSASFIVSVFVLLACVGVANNR1-HIVRSISGPF 164
Db      1 MGVMPHHRNPPKKNLPRIGYSNGSKQVMAIFISQGTIIIVLVLRERVKKFTSPDKRF 60
QY      165 EBAVYAVLVVNIPLPMITPIL-WYEAARKIACLFNDWDFEVLVYQISGHSPLKURQRA 223
Db      61 DEAINVIVPISLLFTNFFLPVASWRGPOVALFKKMWTIVYQKFFKTTGS----- 110
QY      224 VYIAIVPLISVLS---VYITVMTSDLINOVVPYCILDN-----LTAMLGAM 269
Db      111 ---PIVFNLYPLTMSLCVFWLSIAINLSQ---YLOPDRMTWTPAYPIAMLNCF 164
QY      270 ---WFLICEAMSITAHLAERFQKALKHIGPAAVADYRVLMRLSKLTRDTGNALCYTF 326
Db      165 CSIMYINCNAFGTASRALSDALQTTIRGEKPAQKLTREYRHLWVDLSHMMQOUGRAYSNNY 224
QY      327 VEMSLYLFPIITLSIYGLMSQ--LSBEGFGIKDIGITLALMNIIGLFYICDEAHYASVNR 385
Db      225 GMVCLVIFFTTITATYGSIEIIDHGATYKEVGLFVIVCYCGLLYIICNEAHYASRKVG 284
QY      386 TNFOKKLWVEIMNMNSDAQTEINMPLRATENMPTINGCFEPDVNRTLFGKILTTMTY 445
Db      285 LDFQTKLAININTAVDAATQKEVEMLVAINKNPPIIMNDGVANINREILTTNISMTATY 344
QY      446 LVVLLQFOIS 455
Db      345 LVVLLQFKIT 354

```

## RESULT 6

AAE30548 standard; peptide; 33 AA.

```

ID      AAE30548
XX      24-FEB-2003 (first entry)
DT      24-FEB-2003 (first entry)
XX      Fruit fly gustatory receptor peptide, Gr63F1.
XX      Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;
XX      Gr63F1.
XX      Drosophila melanogaster.
XX      WO200268593-A2.
XX      06-SEP-2002.
XX      22-FEB-2002; 2002WO-US005414.
XX      23-FEB-2001; 2001US-0271319P.
XX      (UYCO ) UNITV COLUMBIA NEW YORK.
XX      Axel R, Scott K;
XX      WPI; 2002-698668/75.
XX      Novel nucleic acid encoding insect gustatory or odorant receptor protein
XX      useful for identifying a compound which specifically binds to the
XX      receptor for controlling a pest population in an area.
XX      Disclosure; Page 249; 264pp; English.
XX      The invention relates to an isolated nucleic acid encoding an insect
XX      gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
XX      seven transmembrane domains and a C-terminal domain comprising
XX      consecutive amino acids. The invention is useful for identifying a
XX      compound which activates the insect receptor or inhibits the activity of
XX      the insect receptor. The purified insect receptor protein is embedded in
XX      a lipid bilayer. The invention is sprayed for combating ingestion of
XX      crops by pest insects, combating disease-carrying insects in an area and
XX      controlling a pest population in an area. The invention is useful for

```

CC detecting the presence of insect gustatory or odorant receptor and for  
CC inhibiting the function of the receptor in humans or animals or in  
CC biological fluids isolated from them. The invention is also useful for  
CC identifying or isolating other insect receptors and for combating pest  
CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)  
CC peptide  
XX  
SQ Sequence 33 AA;  
Query Match 6.4%; Score 169; DB 5; Length 33;  
Best Local Similarity 100.0%; Pred. No. 8e-12;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 420 STINGGFDVNRITLFGKLLTMTVTVLVLLQF 452  
DB 1 STINGGFDVNRITLFGKLLTMTVTVLVLLQF 33  
RESULT 7  
AAB75232 ID AAB75232 standard; protein; 433 AA.  
XX AAB75232;  
AC AAB75232;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Drosophila gustatory receptor GR98B.1 protein sequence.  
XX  
KM Gustatory receptor; fruit fly; taste; pheromone; semiochemical;  
XX crop damage; pest control.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200077208-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000MO-US016211.  
XX  
PR 14-JUN-1999; 99US-0138668P.  
XX  
PR 10-FEB-2000; 2000US-0181704P.  
XX  
PA (UYTA ) UNIV YALE.  
XX  
PI Carlsson PJ, Clyne PJ, Warr CG;  
XX  
DR WPI: 2001-061873/07.  
XX  
DR N-PSDB; AAF63771.  
XX  
XX  
PT New isolated nucleic acid molecule encoding Drosophila Gustatory Receptor  
PT protein useful for e.g. identification of compounds which may be used for  
PT pest management.  
XX  
PS Claim 12; Page 214-215; 227pp; English.  
XX  
XX This invention relates to polynucleotide sequences AAF63732 - AAF63777  
CC which encode Drosophila gustatory receptor proteins represented by  
CC sequences AAB75193 - AAB75238. The invention includes methods for  
CC determining gustatory receptor ligands. Also included is a method for  
CC modulating the expression of the DNA encoding the receptors. The DNA and  
CC protein sequences may be used for the identification of compounds, e.g.  
CC pheromones and other semiochemicals, which may be used for pest  
CC management. The DNA sequences may also be used for behavioural studies  
CC involving gustatory systems in various organisms. Also, the DNA sequences  
CC may also be used to track down gustatory receptor genes in insects that  
CC damage crops or transmit diseases  
XX  
SQ Sequence 433 AA;  
Query Match 6.0%; Score 160; DB 4; Length 433;  
Best Local Similarity 20.9%; Pred. No. 6.3e-09;  
Matches 91; Conservative 72; Mismatches 133; Indels 134; Gaps 18;

QY 102 PINMFLRIIGVLPVIRHGPAPAKFEMNSAFISYVVFVLLACVGVYANNR-IHIVRS- 159  
DB 27 PLOPFRTTL-----HKRRGIVILGXACTLISISLWVEECYANIVALQKHKHAHAE 79  
QY 160 ----LSGPF-EBAVIVYVLPVNIIPIMIPILWYBAKIAKLFDNDDEPV-LYYQISG- 212  
DB 80 DSKVMGEVYERSMVMVAMFWNQNLITL-----NFRRLARIYDIDLADLEIDLNASSGF 133  
QY 213 ----HSLPLKLR-QKAVYIAIVY-----PILSVASVITHTWSDILNNOV 253  
DB 134 VGGRHMMRFRFRRLATSGVMIIVLVGTRPFTLVAGPYLHMTNKVTEILIMLQV-KC 192  
QY 254 VPYCIIDNLTMALGAMWFLICBAMSTIAHLAERFQKALKHIGPAAWADRYVLMLELSK 313  
DB 193 TEYCV-----FVLIVYELILKGRHIL-QQISVELE-----GN 223  
QY 314 LTRDTGNALCYTEVFMSLYLPFIITLSIVGLMSQLSGFGIKDIGLTTALMNIIGLLFYI 373  
DB 224 QSRDSVQELC-----VALKKNQQLAGRIMGLVNEVSLYF---TLSLTLFLVYELFTILOI 275  
QY 374 CDEAHYASVY-----VRTYFQKLLMVELMNMNSDAQTEINMFLA--TEMN 418  
DB 276 VNMALIKSVNPNBCCQYSKLVFKRKNTFYKQVIFIIIGVGTCLLSINIFLSCLYSEFC 335  
QY 419 PSTIN-----CGGPFVNRITLFGKL 438  
DB 336 IQTNSISRYVHONYCISAAEDVILMKGLREYSLQMEHLKLTFCGGLDINLKFPGCM 395  
QY 439 LTTMTVTVLVLLQFOI 454  
DB 396 VVTLFGYIIITLVQFKI 411  
RESULT 8  
AAB30509 ID AAB30509 standard; protein; 369 AA.  
XX  
AC AAB30509;  
XX  
DT 24-FEB-2003 (first entry)  
XX  
DE Fruit fly gustatory receptor protein, Gr39D1.  
XX  
KM Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;  
XX Gr39D1.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200268593-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 22-FEB-2002; 2002MO-US005414.  
XX  
PR 23-FEB-2001; 2001US-0271319P.  
XX  
PA (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
PI Axel R, Scott K;  
XX  
DR WPI: 2002-698666/75.  
XX  
XX Novel nucleic acid encoding insect gustatory or odorant receptor protein  
PT useful for identifying a compound which specifically binds to the  
PT receptor for controlling a pest population in an area.  
XX  
XX Disclosure; Page 193-194; 264pp; English.  
XX  
XX The invention relates to an isolated nucleic acid encoding an insect  
CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises  
CC seven transmembrane domains and a C-terminal domain comprising  
CC consecutive amino acids. The invention is useful for identifying a  
CC compound which activates the insect receptor or inhibits the activity of



OS Drosophila melanogaster.  
 XX WO200268593-A2.  
 XX 06-SEP-2002.  
 XX 22-FEB-2002; 2002WO-US005414.  
 XX 23-FEB-2001; 2001US-0271319P.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Axel R, Scott K;  
 XX WPI; 2002-698668/75.  
 XX  
 PT Novel nucleic acid encoding insect gustatory or odorant receptor protein  
 PT useful for identifying a compound which specifically binds to the  
 PT receptor for controlling a pest population in an area.  
 XX  
 XX Claim.3; Page 170-173; 264pp; English.  
 CC The invention relates to an isolated nucleic acid encoding an insect  
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises  
 CC seven transmembrane domains and a C-terminal domain comprising  
 CC consecutive amino acids. The invention is useful for identifying a  
 CC compound which activates the insect receptor or inhibits the activity of  
 CC the insect receptor. The purified insect receptor protein is embedded in  
 CC a lipid bilayer. The invention is sprayed for combating ingestion of  
 CC crops by pest insects, combating disease-carrying insects in an area and  
 CC controlling a pest population in an area. The invention is useful for  
 CC detecting the presence of insect gustatory or odorant receptor and for  
 CC inhibiting the function of the receptor in humans or animals or in  
 CC biological fluids isolated from them. The invention is also useful for  
 CC identifying or isolating other insect receptors and for combating pest  
 CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)  
 CC protein  
 XX  
 XX Sequence 736 AA;  
 SQ  
 Query Match 5.0%; Score 132.5; DB 5; Length 736;  
 Best Local Similarity 21.0%; Pred. No. 2.6e-05;  
 Matches 77; Conservative 66; Mismatches 134; Indels 89; Gaps 16;  
 QY 102 PIMFLRIIGVLPIVRHGPAPAKFEMNSASFIYSVFVLLACVGVANNR-IHIVRS- 159  
 DB 427 PIGFRTLT-----HKRRRGIVILGVACVLISISLWVIECYANIVALKDIIHKFHA 479  
 QY 160 ----LSGPFEBAYIAYFLVNIPLPIMIIPIIMYEARKIAKLFNDWDPEV-LYYQISG-- 212  
 DB 480 DSKVMGNTQKVAVVAFVFNQNLIL-----NFRRLARIYDIDALEIDLNNASSGEV 533  
 QY 213 ---HSLPLKLR-QAVVYIAYL-----PILSLSVITHTVMSDINNOV 254  
 DB 534 GQRHMFRRRLALSVGLMIVLVGLPRFTVLVAGPLHTMTNVILHILIMQL-KCT 592  
 QY 255 PYCILDNLTLTAMLGAWFLICBAMSTIHLAERFOKALKHIGPAMVADYVLMRLSKL 314  
 DB 593 EYCV-----FVLLIYELIRGRHIL-QQISVELE-----GNQ 623  
 QY 315 TRDGNALCTFVMSLYLFTITLSTYGLMSQISEGGIDIGLITLALNIGLPTIC 374  
 DB 624 SRSSVQELC-----VALKRQQLAGRIWGLVNEVSLVF--TSLTLTFLTNELTTIQIV 675  
 QY 375 DEAHYASVNR--TNFOKLLMVEELNMMNSDAQTEINMFLRATENPSTINCSPFDVNR 432  
 DB 676 NMALIKSVNPECOCYEDYILIKGLAREVSLQWHEHLKLI-----FTCCGLFDINTL 726  
 QY 433 TLFKGL 438  
 DB 727 KEFGGV 732

RESULT 11  
 ABB70025  
 ID ABB70025 standard; protein; 1299 AA.  
 XX  
 AC ABB70025;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 36867.  
 XX  
 KM Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL14128.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 36867; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1299 AA;  
 Query Match 5.0%; Score 132.5; DB 4; Length 1299;  
 Best Local Similarity 19.3%; Pred. No. 6.6e-05;  
 Matches 81; Conservative 75; Mismatches 159; Indels 105; Gaps 19;  
 QY 96 FYRNIDPIMFLRIIGVLPV--VRH-GPAPAKFEMNSASFIYSVFVLLACVGVANN 152  
 DB 484 FHRVSNVLFISQYGLPVSNNVADVDVDRFMCSPRIIYSLIGLINSRGAVINY 543  
 QY 153 RHIVRSLSGPFEBAYIAYFLVNIPLPIMIIPIIMYEARKIAKLFNDWDPEVLYYQISG 212  
 DB 544 VIKTIN-----FHSSTLSLYVCLLEHL--FWRLAIQWPRIMRTWVGVEQLFLRPVY 596  
 QY 213 HSLPLKLRQKAVYIAYVPLISV-----SVTITHVMSDLNIN----- 251  
 DB 597 RFYGEYRIKRIYIVFTIVMSALVEHCLLIGNSFHLNMRTOCKINVTYFESIYKWR 656  
 QY 252 ----QVVPYCLDNLTLTAMLGAW-----W-----PLICEAMSTI--HLLA 285  
 DB 657 PHLYMILPYH-----WMLPILEWVNGTIAIRSPFDCRIMCIGICLAARFHO 706  
 QY 286 ERFOKALKHIGPAMVADYVLMRLSKLTR--DTGNALCTYFV--MSIYLFPIITLS 340  
 DB 707 RRIAAVHRKWPVAVFWTEVREHYLALKRLVHLDDAALPVLVLAFGNMSPICF----- 760



QY 341 IYGLMSQSEGGIKDIDG---LTITLW-NIG-----LFFICEAHYAVNVRNTO 389  
 DB 761 -----QLFNSF--KNIGDFLWMLAFWYSLGFAVVRTLLITFVA-----SSIN--DYE 804  
 QY 390 KKLMLVLMNMSDAQTEINMFLRATENMPSSTINCSPFDVNRTEFKGLITTMVTVL 449  
 DB 805 RKIVTALRDVPSRAMSIEVGRFSEQJGNDTALSGSGFFYLFRSLVLMAGTIIITFELMI 864

RESULT 12  
 ID AAB75208 standard; protein; 372 AA.  
 AC AAB75208;  
 DT 03-APR-2001 (first entry)  
 DE Drosophila gustatory receptor GR39D.2a protein sequence.  
 KM Gustatory receptor; fruit fly; taste; pheromone; semiochemical;  
 KW crop damage; pest control.  
 OS Drosophila melanogaster.  
 PN WO20077208-A2.  
 PD 21-DEC-2000.  
 PF 14-JUN-2000; 2000WO-US016211.  
 PR 14-JUN-1999; 99US-0138668P.  
 PX 10-FEB-2000; 2000US-0181704P.  
 PA (UYVA ) UNIV YALE.  
 PI Carlson PJ, Clyne PJ, Warr CG;  
 DR WPI; 2001-061873/07.  
 N-PSDB; AAF63747.  
 PT New isolated nucleic acid molecule encoding Drosophila Gustatory Receptor  
 protein useful for e.g. identification of compounds which may be used for  
 pest management.  
 PS Claim 12; Page 129-130; 227pp; English.  
 XX This invention relates to polynucleotide sequences AAF63732 - AAF63777  
 CC which encode Drosophila gustatory receptor proteins represented by  
 CC sequences AAB75193 - AAB75238. The invention includes methods for  
 CC determining gustatory receptor ligands. Also included is a method for  
 CC modulating the expression of the DNA encoding the receptors. The DNA and  
 CC protein sequences may be used for the identification of compounds, e.g.  
 CC pheromones and other semiochemicals, which may be used for pest  
 CC management. The DNA sequences may also be used for behavioral studies  
 CC involving gustatory systems in various organisms. Also, the DNA sequences  
 CC may also be used to track down gustatory receptor genes in insects that  
 CC damage crops or transmit diseases  
 XX  
 SQ Sequence 372 AA;  
 Query Match 4.9%; Score 130.5; DB 4; Length 372;  
 Best Local Similarity 21.7%; Pred. No. 1.5e-05;  
 Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;  
 QY 129 SASFSYVVFLLACTYGVYVANNRIHL-VRSLSG-PFEENY-----AYFL--- 174  
 DB 39 SSTAQIVVGVFMALGALLESIVMETKSQGTGTFPNNAVILTTSVTQLANMLRSQ 98  
 QY 175 ----VNILPIM--IIPILMYEARKIAKLFWNDDEFLVYQISGSLPLKAKQKAVYIAI 228  
 DB 99 QKSQVNLQRLSQVVEVLQFEPYAVFO-----FRMLY-----RI 132

QY 229 VPILSVLSVITHTVMSDLINOV-----VPYCIIDLNTAMLGAMFELICEAMSI 279  
 DB 133 MLVVCILYGAAMTHREINMLTMOISRVLTILGFVRCILAN-----FOFTCYGMV 185  
 QY 280 TNLHAEFPQALKHI-----GPAAVADY-RVLMRLSKLRTDGNALCYTFVF- 328  
 DB 186 ILKKLQVQVKOLEHLVSTTISMGVAGCLRTHEBILLLGORELTAVVGAVLLFLFIYQ 245  
 QY 329 --MSLVFFIITLSIGMSQSEGGIKDIDGTLTALNIGLFFICEAHYAVNVRN 386  
 DB 246 VMQCILIFITISNL-----EGFHSN-DVLIFCMLAPMLFYLLIPLVYNDIHQA 294  
 QY 387 NFQKLMVLMNMSDAQTEINMFLRATENMPSSTINCSPFDVNR-TLFKGLITTMVTV 445  
 DB 295 NKTAKVL-TKVPRTGTGLDRMIKFLKRLKQKPLITAGFPAIDKSTLFK-LFTAIPTY 352

QY 446 LVVLIQFO 453  
 DB 353 MWLVQFK 360

RESULT 13  
 ID AAE30510 standard; protein; 372 AA.  
 AC AAE30510;  
 DT 24-FEB-2003 (first entry)  
 DE Fruit fly gustatory receptor protein, Gr39D2a.  
 KM Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;  
 KW Gr39D2a.  
 XX Drosophila melanogaster.  
 OS  
 PN WO200268593-A2.  
 PD 06-SEP-2002.  
 PF 22-FEB-2002; 2002WO-US005414.  
 PR 23-FEB-2001; 2001US-0271319P.  
 PX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PI Axel R, Scott K;  
 DR WPI; 2002-698668/75.  
 XX Novel nucleic acid encoding insect gustatory or odorant receptor protein  
 CC useful for identifying a compound which specifically binds to the  
 CC receptor for controlling a pest population in an area.  
 PS Disclosure; Page 194-196; 264pp; English.  
 XX The invention relates to an isolated nucleic acid encoding an insect  
 CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises  
 CC seven transmembrane domains and a C-terminal domain comprising  
 CC consecutive amino acids. The invention is useful for identifying a  
 CC compound which activates the insect receptor or inhibits the activity of  
 CC the insect receptor. The purified insect receptor protein is embedded in  
 CC a lipid bilayer. The invention is sprayed for combating ingestion of  
 CC crops by pest insects, combating disease-carrying insects in an area and  
 CC controlling a pest population in an area. The invention is useful for  
 CC detecting the presence of insect gustatory or odorant receptor and for  
 CC inhibiting the function of the receptor in humans or animals or in  
 CC biological fluids isolated from them. The invention is also useful for  
 CC identifying or isolating other insect receptors and for combating pest  
 CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)  
 CC protein  
 XX  
 SQ Sequence 372 AA;



Query Match 4.9%; Score 130.5; DB 5; Length 372;  
 Best Local Similarity 21.7%; Pred. No. 1.5e-05;  
 Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;

QY 129 SASFIYVVPFVLACVGVANNRHHI-VRSLSG-PPEBAVI-----AYLPL- 174  
 DB 39 SSTAIGIVVGVFMALLGALASLYYMETKSGTGFDAVILITTSVTOLLANMLRSQ 98  
 QY 175 ----VNILPIM--IIPILWYEAARKIALFNDWDFEVLYYQISGSLPLKLRQKAVYIAI 228  
 DB 99 QKSOVILLQRLSQVVELLOPEYAVPQ-----FRLLY-----RI 132  
 QY 229 VLPILSVLVVITHVTSNDINQV-----VPCIIDNLTAMLGAMWFLICEAMST 279  
 DB 133 WLIVCLIGYAMVTHFGINMLTQISRVLTIGFVYKCVLAN-----FOFCYGTGMV 185  
 QY 280 TAILLAERPOKALKHI-----GPAWAVDY-RVLMRLSKLTRDGNALCYTFVF- 328  
 DB 186 ILKKLQOVQKOLEHLVSTTTTISMAVGAGCLRTHDEILLAGORELAVVGGVILFLFIYQ 245  
 QY 329 --MSLVFFITITSIYGLMSQLSEGFQIKDIGLITLANNIGLFYICDEAHYASVAVRT 386  
 DB 246 VMOCILFIYSNL-----EGFHSN-DLVILICWLAPEMLFYILPLVVDINHQA 294  
 QY 387 NPOKRLMVELMNMNSDAQTEINMFLRATENMPSSTINGCFPDVNR-TLEKGLITTVTY 445  
 DB 295 NTKAKML-TKVPRTGTLDNRHIEKFLKRLKQKILTAVGFPALDKSTLEK-LFTALFTY 352  
 QY 446 LVVLLQFQ 453  
 DB 353 MVLVQPK 360

RESULT 14  
 ABB68743  
 ID ABB68743 standard; protein; 436 AA.  
 AC ABB68743;  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 33021.  
 XX  
 KM Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN W0200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL12846.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 33021; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pot\_sequences  
 XX  
 SQ Sequence 436 AA;  
 QY  
 DB Query Match 4.9%; Score 129; DB 4; Length 436;  
 Best Local Similarity 21.2%; Pred. No. 2.9e-05;  
 Matches 94; Conservative 70; Mismatches 172; Indels 108; Gaps 22;

QY 85 ASNNPEKP-SVFRYNDIPNMFRLIIGVPL---YRHGKRAKFEKNSAFIYVVFV 140  
 DB 30 AQGGLFEQDFTFYGAIRPYLCVAQFEGIMPLSNIRSDQDVKFKRSIGLAVTGFL 89  
 QY 141 L--LACVGVANNRHHIYVRSLSGPPEBAVIAYLVNLPIMIPILWYE-----ARKIA 194  
 DB 90 LGGKKTIVG--AN-----ILFTEGLNAKNIVGLVFLI-----VGMVNLNPFGRASWS 136  
 QY 195 KLFNDWDFEVLV-----YQISGSLPLKLRQKAVYIAIVLPILSV----- 235  
 DB 137 HMLPMSVDIIMLPFYKRGKRS-----LRSKNVIALSVVLAAGDHMLYASGYCSYS 192  
 QY 236 LSVVITHVTSNDL-----NINQVVPYCILDNLTM-LGAW-----WFLICE 275  
 DB 193 MHILQCHTNSRIITFGLYLEKFSDFIMPF-----NIFMCYGFWMNGAFTEPLNFMDI 248  
 QY 276 AMSTTALLAERPOKALKHI-G-----PAWAVDYRVLMRLSKLTRDGNALCYTFVF 328  
 DB 249 FIWTSIGLAQRFQGFARVGALEGRHVPALWDIRDHRLCELA-----SLVEAS 301  
 QY 329 MSVLFFITITSIYGLMSQLSEGFQIKDIGLITLALM-----NIGLIFYICDEAHYA 380  
 DB 302 MSNIVFVSCANNVYVIGNQALAITKLRHPINYYFYSLIFLARTSLVFMASKIHDA 361  
 QY 381 SVNVRNPOKRLMVELMNMNSDAQT-EINMFLRATENMPSSTINGCF--FDVVRITLPGK 437  
 DB 362 SL-----LPLRSYLVPDSQDQVORF--ADQLTSEFVGLSGYRFLCTRKSLFG 410  
 QY 438 LITTMVTVLVLLQFOISIFPDG 461  
 DB 411 MLATLVTEMLL--QIDAKSHKG 432

RESULT 15  
 AAE30491  
 ID AAE30491 standard; protein; 436 AA.  
 AC AAE30491;  
 XX  
 DT 24-FEB-2003 (first entry)  
 DE Fruit fly gustatory receptor protein, Gr61D1.  
 XX  
 KM Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;  
 KM Gr61D1.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN W0200268593-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-US005414.  
 PR 23-FEB-2001; 2001US-0271319P.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2005, 19:15:31 ; Search time 43 Seconds  
(Without alignments) 888.845 Million cell updates/sec

Title: US-10-081-816-12

Perfect score: 2645  
Sequence: 1 MKPSEKRVKVGHGQNSGHS.....TTTGTTLPPIMKLGKRGK 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	4.1	106	US-09-270-767-42643	Sequence 42643, A
2	107	4.0	404	US-09-270-767-33959	Sequence 33959, A
3	107	4.0	404	US-09-270-767-49176	Sequence 49176, A
4	104.5	4.0	323	US-09-134-001C-5200	Sequence 5200, Ap
5	104	3.9	363	US-09-491-577-74	Sequence 74, Appl
6	102	3.9	907	US-09-198-452A-306	Sequence 306, App
7	102	3.9	928	US-09-438-185A-295	Sequence 295, App
8	101	3.8	239	US-09-710-279-1696	Sequence 1696, Ap
9	101	3.8	395	US-09-491-577-98	Sequence 98, Appl
10	100.5	3.8	518	US-09-134-001C-4744	Sequence 4744, Ap
11	99.5	3.8	1989	US-09-949-016-10076	Sequence 10076, A
12	99	3.7	1900	US-09-538-092-564	Sequence 8212, Ap
13	98.5	3.7	656	US-09-489-039A-8212	Sequence 6, Appl
14	98	3.7	360	US-08-671-525B-6	Sequence 6, Appl
15	98	3.7	360	US-08-672-109B-6	Sequence 6, Appl
16	98	3.7	360	US-08-842-045-6	Sequence 6, Appl
17	98	3.7	360	US-08-842-238-6	Sequence 6, Appl
18	98	3.7	360	US-08-780-749A-1	Sequence 1, Appl
19	98	3.7	360	US-08-629-335B-6	Sequence 6, Appl
20	98	3.7	360	US-08-870-511-1	Sequence 1, Appl
21	98	3.7	360	US-09-709-066-4	Sequence 4, Appl
22	98	3.7	600	US-09-134-000C-5694	Sequence 5694, Ap
23	97.5	3.7	459	US-09-097-889-22	Sequence 22, Appl
24	97.5	3.7	459	US-09-098-079-22	Sequence 22, Appl
25	97.5	3.7	486	US-09-134-001C-3593	Sequence 3593, Ap
26	97	3.7	309	US-09-328-352-7637	Sequence 7637, Ap
27	97	3.7	391	US-07-816-283-2	Sequence 2, Appl

28	97	3.7	391	US-08-417-103-2	Sequence 2, Appl
29	97	3.7	391	US-08-417-103-14	Sequence 14, Appl
30	97	3.7	391	US-09-826-509-569	Sequence 569, App
31	96.5	3.6	450	US-09-721-870-16	Sequence 16, Appl
32	96.5	3.6	551	US-10-327-189-7	Sequence 7, Appl
33	96.5	3.6	551	US-09-521-195B-1	Sequence 1, Appl
34	96.5	3.6	822	US-09-824-734-3	Sequence 3, Appl
35	96.5	3.6	808	US-09-134-001C-3105	Sequence 3105, Ap
36	95.5	3.6	209	US-09-248-796A-21116	Sequence 21116, A
37	95	3.6	391	US-08-120-601B-8	Sequence 8, Appl
38	95	3.6	436	US-09-543-681A-4395	Sequence 4395, Ap
39	94.5	3.6	1116	US-09-543-681A-4379	Sequence 4379, Ap
40	94	3.6	411	US-09-489-039A-7912	Sequence 7912, Ap
41	94	3.6	520	US-09-252-991A-17058	Sequence 17058, A
42	93.5	3.5	551	US-10-327-189-38	Sequence 38, Appl
43	93	3.5	391	US-07-816-283-4	Sequence 4, Appl
44	93	3.5	391	US-08-417-103-4	Sequence 4, Appl
45	93	3.5	1058	US-09-328-352-4276	Sequence 4276, Ap

#### ALIGNMENTS

```
RESULT 1
US-09-270-767-42643
Sequence 42643, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42643
LENGTH: 106
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-42643

Query Match 4.1%, Score 109, DB 4, Length 106,
Best Local Similarity 45.1%, Pred. No. 0.0012,
Matches 23, Conservative 9, Mismatches 19, Indels 0, Gaps 0;

Cy 405 QTEINMFLRATKEMNPSTINGCPFDVRRTPFKGLITMTYTVLVLPQIS 455
Db 46 QKEVEMLVATIKNPIMNLDGVANINRELIITWISPMATYLVLLQFKIT 96

RESULT 2
US-09-270-767-33959
Sequence 33959, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 33959
LENGTH: 404
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-33959

Query Match 4.0%, Score 107, DB 4, Length 404,
Best Local Similarity 18.2%, Pred. No. 0.016,
Matches 75, Conservative 85, Mismatches 155, Indels 96, Gaps 17;

Cy 106 FLRIIGVLPYRHGPABAKFEMNSAFIYSVVFVLLACVGVYANNRIHVRSLSGPFE 165
```

```
Db 23 FMTVFGLLANRYRAGRERFRPSKANTLAPASLMAINFSLVYG-----RQIYKEQ 72
QY 166 E-----AVIAYLFL-----VNILPIMILPMTBAKRIAKLFDNDMDFEVLTYQISG 212
Db 73 EQQINLKQATLTYSYVNIIVAVINVSQMI-----SDHVAKVLSKVFPFDTL----- 120
QY 213 HSLPLKROKAVYIAVLPILSVSVIT-----HVTMS-----DNTIN 251
Db 121 --KERLDRSRSLYSIVLALVKTVAEPFLITVAFLIQGRORQRPBMSLWTLVRLPFLIIS 178
QY 232 QVVPYCIIDNLTAMLGAMWFL--ICEAMSITAHILAERFOKAL--KHIGPAAWAD-- 303
Db 179 NFLNNCYFGAMVYVKEILYALNRRLERAOEVLNLOKQOLKLYTYKYMORFCALADBL 238
QY 304 -----YRVLMLRLSKLTRDTG-----NALCY-----TFVPMSLYLFITLISYGLMSQL 348
Db 239 DQLAVRYRLIYVHSGCYLTPMSLSMILSLICHLGITVGFSLYYAIAIDL-----IMGRP 294
QY 349 SEGFG--IKDIGITITLWNLGILFYICDEAHYASVNVRTNFOKLLMELMMNSDA-- 404
Db 295 YDGLSLLNLVPLSI-SLAETITLTLCHNLVAT-----RSAAVILQEMALQHADSRX 347
QY 405 QTEINMFLATEMNPSTINGCGFEDVNRTLFKGLLTTWTVYLVLLOFOIS 455
Db 348 RQAVHGFLLVTVTKYQIKPLGLYELDMRLISNVFSAVASFLILIVQADLS 398
```

RESULT 3  
US-09-270-767-49176  
Sequence 49176, Application US/09270767  
Patent No. 6703491

GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 49176  
LENGTH: 404  
TYPE: PRF  
ORGANISM: Drosophila melanogaster  
US-09-270-767-49176

Query Match 4.0%; Score 107; DB 4; Length 404;

Best Local Similarity 18.2%; Pred. No. 0.016;  
Matches 75; Conservative 85; Mismatches 155; Indels 96; Gaps 17;

```
QY 106 FLRIIGVLPYVHGPAPAKFEMNSAFIYSVVFVLLACYGVVANNRIHIVRSLSGPR 165
Db 23 FMTVFGLLANRYRAGRERFRPSKANTLAPASLMAINFSLVYG-----RQIYKEQ 72
QY 166 E-----AVIAYLFL-----VNILPIMILPMTBAKRIAKLFDNDMDFEVLTYQISG 212
Db 73 EQQINLKQATLTYSYVNIIVAVINVSQMI-----SDHVAKVLSKVFPFDTL----- 120
QY 213 HSLPLKROKAVYIAVLPILSVSVIT-----HVTMS-----DNTIN 251
Db 121 --KERLDRSRSLYSIVLALVKTVAEPFLITVAFLIQGRORQRPBMSLWTLVRLPFLIIS 178
QY 232 QVVPYCIIDNLTAMLGAMWFL--ICEAMSITAHILAERFOKAL--KHIGPAAWAD-- 303
Db 179 NFLNNCYFGAMVYVKEILYALNRRLERAOEVLNLOKQOLKLYTYKYMORFCALADBL 238
QY 304 -----YRVLMLRLSKLTRDTG-----NALCY-----TFVPMSLYLFITLISYGLMSQL 348
Db 239 DQLAVRYRLIYVHSGCYLTPMSLSMILSLICHLGITVGFSLYYAIAIDL-----IMGRP 294
QY 349 SEGFG--IKDIGITITLWNLGILFYICDEAHYASVNVRTNFOKLLMELMMNSDA-- 404
Db 295 YDGLSLLNLVPLSI-SLAETITLTLCHNLVAT-----RSAAVILQEMALQHADSRX 347
```

```
QY 405 QTEINMFLATEMNPSTINGCGFEDVNRTLFKGLLTTWTVYLVLLOFOIS 455
Db 348 RQAVHGFLLVTVTKYQIKPLGLYELDMRLISNVFSAVASFLILIVQADLS 398
```

RESULT 4  
US-09-134-001C-5200  
Sequence 5200, Application US/09134001C  
Patent No. 6380370

GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5200  
LENGTH: 323  
TYPE: PRF  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5200

Query Match 4.0%; Score 104.5; DB 3; Length 323;  
Best Local Similarity 22.7%; Pred. No. 0.02;  
Matches 58; Conservative 34; Mismatches 79; Indels 85; Gaps 10;

```
QY 271 FLICEAMSTAHILAEERFOKALKHIGPAAWADY-RVLMLRLSKLTRDTGNALCYTFVFM 329
Db 66 FLINDBEAQVHLSNFEFMSNMHL--LGTDYGRDLFSRLVSGSRAT-----LFTVL 116
QY 330 SLYPFIITLSYGLMSQISEGF-----GINDIGITI-----TALNIGL 369
Db 117 LTLFTVVGVPGLLAGYKKGMIDTIIRIRIDIGLSIDEPYIMIALASFHPSLWNLVI 176
QY 370 LFYICDEAHYASVNVRTNFOKLLMELMMNSDAQTEINMFLATEMNPSTINGCFPD 429
Db 177 ATTI-----IKMN--YTRVTRGIVTEMNQSYIQMAQFEN 210
QY 430 VN--RTLFGKLLTTWTVYLVLLOFOISIPDTGDSGANNTIVDFVNDSDNDMSLWG 487
Db 211 VSTNAILFRLHLPKXLPBSFVIM-----IYDF-----GKIILY 243
QY 488 ASTLSTTVGTTLPPP 503
Db 244 ISSLSFLGLGQPPSP 259
```

RESULT 5  
US-09-491-577-74  
Sequence 74, Application US/09491577  
Patent No. 6610511

GENERAL INFORMATION:  
APPLICANT: Yale University  
APPLICANT: Carlson, John R.  
APPLICANT: Kim, Hunhyong  
APPLICANT: Clyne, Peter J.  
APPLICANT: Watt, Coral G.  
TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila  
FILE REFERENCE: 44574-5061-US  
CURRENT APPLICATION NUMBER: US/09/491,577  
CURRENT FILING DATE: 2000-01-25  
EARLIER APPLICATION NUMBER: US 60/117,132  
EARLIER FILING DATE: 1999-01-25  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 74  
LENGTH: 363

```

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-491-577-74

Query Match      3.9%; Score 104; DB 4; Length 363;
Best Local Similarity 22.9%; Pred. No. 0.028;
Matches 77; Conservative 43; Mismatches 124; Indels 92; Gaps 17;

Qy 122 RAKEKNSASF--TYSVFVLLACTGVYANNRHHVSLSGPFEAVIAYLP-LVNI 177
Db 109 RSKQEVMMREHRRFNRVFMFYCISAGVI-----PF--IVIQPLFDLPNR 153
Qy 178 LPIWIIPIIWEARKIKALFNDMDDFEVLXXYQISGHSIPLKROKAVIAYLPILSTLS 237
Db 154 LPPV---MWTPF-----DWQOPVLLMYAP-----IYQATITPL----- 183
Qy 238 VITHVTSMDINQVVPYCIIDNLTMGLAMWFLICEAMSTHLLAEFOKALKHIGP 297
Db 184 ACAGNVMTDAVN-----WYIMLH-LSLCLRMGLGRSLK-LQHDD- 220
Qy 298 AAMVADYVLMRLSKLTRDTGNALCTYFVMSLYLPITLISYGLMSQISEGFIKDI 357
Db 221 ---KDIKEKLEIHLHQRLKQALSIEIFISKSTFQI-----LVSSILICFTIYSM 270
Qy 358 GLTTTALNIGLLFYICDEAHYASVNRVTNFOKKLAWELMW--MNSDAQEINMFLBAT 415
Db 271 QMYIVAMIMQVMLPTTGNNAVIDSANMLTD-----SMTNSDMPDKNCRRRLVLMFM--V 323
Qy 416 EMN-PTSTINGCFEVDNRTLFKGLTTWVTVLVVL 450
Db 324 YLNPVTLKAGFFHIGLPLFTKTMNQAYSLALLL 359

RESULT 6
US-09-198-452A-306
; Sequence 306, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof, and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 306
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-306

Query Match      3.9%; Score 102; DB 4; Length 907;
Best Local Similarity 18.8%; Pred. No. 0.18;
Matches 90; Conservative 71; Mismatches 149; Indels 168; Gaps 25;

Qy 88 NPEKPSVRYNIDPIMFLRIIGVLPVRHGPARAKEKNSAS--FYSVFVLLAC 144
Db 112 DPPSPLEFFYR---IVMSLTILSYTSF--MGFVDQFNLQDGKHFICFNAILIFLGAI 166
Qy 145 YGVVANNRHHVSLSGPFEAVIAYLPVNIILPIMIIPILMYEARKIKALFNDMDPE 204
Db 167 GSGIIL-SLVHTI-----GIQGIILFTALVLPFIVFYVSKLSKLSDDHD--- 213
Qy 205 VLYYQISGHSIPLKROKAVIAYLPILSVLSVITHVTSMDINQVVPYCIIDNLTA 264
Db 214 -LFLDTHHPPL-----SKALKCFYDKTYF 238
Qy 265 MGLAMWFLICEAMSTTA---HLLAEFOKALK-----HIGPAMVADYVLMRL--- 311
Db 239 YLTCFYPLM-QLALATFENYLIKFEIOPASKEBEFELVAHIGKCS-----LWISLGNM 290
Qy 312 -----SKLTRDNG--NALCYT-FVFNLSLYF--FIITLSIYGLMSQISEG----- 351

```

```

Db 291 CFALFAYSRIYKRGVNNIILFAPLCTLSLFLFWTFKTTISIAVLAMVREGVYALDDN 350
Qy 352 ---FG-----IKDIGTITLALMN-----IGLL--FY 372
Db 351 NLAGIIVGPNKIRNQRIIVESFIEBIGMLV---WSLVCFSSQOYVFCILISIAITL 407
Qy 373 ICDEAHYASVNRVTNFOKKLAWE--LNMNS--DAQEINMFLRATENMPSITNCGG 426
Db 408 VCLVRSYAKAILNLSAQLQTRSMQDWIKSMTVKOKROVELPLLAHLKHPSEBROT 467
Qy 427 FFD-----VNRTEFKGLTTWVTVLVVLQFOISIPTDKDGSEGANNTVVDPYWDSL 479
Db 468 AFQHLNLNASRSLVPSILAHN-----NKLSLP-----NKLKTIEMVSKSL 507

RESULT 7
US-09-438-185A-295
; Sequence 295, Application US/09438185A
; Patent No. 682071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kaiman, Sue
; APPLICANT: Davis, Ronald
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 295
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0293
US-09-438-185A-295

Query Match      3.9%; Score 102; DB 4; Length 928;
Best Local Similarity 18.8%; Pred. No. 0.13;
Matches 90; Conservative 71; Mismatches 149; Indels 168; Gaps 25;

Qy 88 NPEKPSVRYNIDPIMFLRIIGVLPVRHGPARAKEKNSAS--FYSVFVLLAC 144
Db 112 DPPSPLEFFYR---IVMSLTILSYTSF--MGFVDQFNLQDGKHFICFNAILIFLGAI 166
Qy 145 YGVVANNRHHVSLSGPFEAVIAYLPVNIILPIMIIPILMYEARKIKALFNDMDPE 204
Db 167 GSGIIL-SLVHTI-----GIQGIILFTALVLPFIVFYVSKLSKLSDDHD--- 213
Qy 205 VLYYQISGHSIPLKROKAVIAYLPILSVLSVITHVTSMDINQVVPYCIIDNLTA 264
Db 214 -LFLDTHHPPL-----SKALKCFYDKTYF 238
Qy 265 MGLAMWFLICEAMSTTA---HLLAEFOKALK-----HIGPAMVADYVLMRL--- 311
Db 291 CFALFAYSRIYKRGVNNIILFAPLCTLSLFLFWTFKTTISIAVLAMVREGVYALDDN 350
Qy 352 ---FG-----IKDIGTITLALMN-----IGLL--FY 372
Db 351 NLAGIIVGPNKIRNQRIIVESFIEBIGMLV---WSLVCFSSQOYVFCILISIAITL 407
Qy 373 ICDEAHYASVNRVTNFOKKLAWE--LNMNS--DAQEINMFLRATENMPSITNCGG 426

```

Db 408 VCLVSVYAKAILKLNLSAQLQTRSMQWIKSMYKQKQVELFLAHKHPSEHQF 467  
Qy 427 FPD-----VNRFLFKELTMTTYLVVLLQFOISITDKDSEGANNTVDFVMSL 479  
Db 468 AFQHLNLINLASRSLPSILAHM-----NKLSTP-----NKLKTIEMVXSSL 507

## RESULT 8

US-09-710-279-1696  
; Sequence 1696, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1696  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-1696

Query Match 3.8%; Score 101; DB 4; Length 239;  
Best Local Similarity 23.7%; Pred. No. 0.03; Mismatches 61; Indels 82; Gaps 9;  
Matches 53; Conservative 28; Mismatches 61; Indels 82; Gaps 9;  
Qy 303 DY-RVIMLRSLKTRDTGNALCTYFVMSLYLFFITLSIYGLMSQLSHG-----GIK 355  
Db 11 DYGRDLFSRLVVGSRAT-----LFTVLTLLFTVVGVPLGLLAGYKKGKMDITIMRII 64  
Qy 356 DIGLIT-----TALMNLGLFYICDEAHYAVVVRNTPQKKLMTVELMNN 401  
Db 65 DIGLSLPEFVIMTALASFPHPSLMNVIAIT-----IKMNN 101  
Qy 402 SDQTEINNFPLATEMNPSTINGCFEDVNA--RTLFKGLTMTVTVLVVLLQFOISIPFD 459  
Db 102 ---YRVRTRGIWVTEWNGYIQAGFFNVSTINILPKHLIPKVLPSIFVIM----- 149  
Qy 460 KQDSGANNITVDFVMSLNDMSIMGASTLSTTTVGTTLPPP 503  
Db 150 -----IVDF-----GKILYISLSFLGIGAQPSP 175

## RESULT 9

US-09-491-577-98  
; Sequence 98, Application US/09491577  
; Patent No. 6610511  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; APPLICANT: Carlsen, John R.  
; APPLICANT: Kim, Hanyong  
; APPLICANT: Clyne, Peter J.  
; APPLICANT: Warr, Coral G.  
; TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila  
; FILE REFERENCE: 44574-5061-US  
; CURRENT APPLICATION NUMBER: US/09/491,577  
; CURRENT FILING DATE: 2000-01-25  
; EARLIER APPLICATION NUMBER: US 60/117,132  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 98  
; LENGTH: 395  
; TYPE: PRT

; ORGANISM: Drosophila melanogaster  
US-09-491-577-98

Query Match 3.8%; Score 101; DB 4; Length 395;  
Best Local Similarity 18.9%; Pred. No. 0.066;  
Matches 56; Conservative 58; Mismatches 112; Indels 70; Gaps 12;  
Qy 222 KAVYIAIVL---PILSV-----LSVYTH-----VTMSDNLNVQVY-- 256  
Db 97 KSFYIELSDDPITINILVKETTRLSVLSRINLMSGCTCGFTYTPFSGSERVLPGM 156  
Qy 257 -----CILDNLTMGAMWFLICEANSTPAHLAERFQALRH-IGP 297  
Db 157 VLPTEYKXASPYEIEFVIOAIMAPMGCCMYIPYTNVVFVFLPALIMCVALQKLS 216  
Qy 298 AAMVADRY-----LW-----LRLSKTRTGNALCTYFVMSLYLF--IITLSIYGLMS 346  
Db 217 LRLKNEQVRGEIWCIRYQKLKSGFV--DSMNL-NTHLHVEFLCFGAMLCVLLPSLIT 274  
Qy 347 QISEGFGIDGLTTALMNLGLFYICDEAHYA-----SVVRNTPQKKLM 394  
Db 275 AQTIAQTVIVIAMWMIIPANSVLYVANLEYQVRVQSFRLKYLKGIQSGFDALIA 334  
Qy 395 VELANMNSDAQTEINM-FLRATENPSTINGCFEDVNRTEFLKGLTMTVTVLV 449  
Db 335 YESNMWDFVDVDTQKTLKFLMRSQKPLATLVGGTVPMNLKMLQSLINAIYSFTLL 390

## RESULT 10

US-09-134-001C-4744  
; Sequence 4744, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4744  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4744

Query Match 3.8%; Score 100.5; DB 3; Length 518;  
Best Local Similarity 19.9%; Pred. No. 0.11;  
Matches 94; Conservative 63; Mismatches 139; Indels 177; Gaps 21;  
Qy 93 PSVFYRNIDPIMFLRIIGVLPVHGRPAKFEWNSAFIYSVVFVLLACTYGVYANN 152  
Db 130 PDFFKRLDDKKNIKIISGLII-----VVFETLYT-HSGFVSGG 168  
Qy 153 RHIVRSLSGPFEE-----VAYLEFLVNIILP-----IMIPIIM 187  
Db 169 KL-----FESAPGLNTHAGLLIIVILVFYTFGGYIAVSTIDPFQGVIMLAWM 219  
Qy 188 YEARKIAKLFDNDPFE-----VLYYQ 209  
Db 220 VPIVALIKL-KGNDTFPHDAQMKPTNLDLFRGTVLGVISLSKGLGYRGQHIIVRFS 278  
Qy 210 IGHSLPLKLRQKAV-YIAIVLPILSVLSVIT-----HTWSD-----LNIQVVP 255  
Db 279 IKSHKLLPKARLGLSMAVGL--LGAIGVGLTGISFISERHAKLEDPEPTLPIVMSQILF 336  
Qy 256 YCILDN--LTMALGAMWFLICAMSTIATALLAERFQALKHIGPAAWADVIVLMLRSLK 313  
Db 337 HPLVGGFLAAITLAIMSTISSQLVITSSSLTEDEYKILIRGSDKASHQKEFVLIGRLSV 396

QY 314 LTRDGNALCYTVFVNSLPIFIITL-----SIYGLMSQLSGFGIKDGLTTTALMNI 367  
 Db 397 L-----LVAIVAITAMHPDITLNLVGNAMGFGAASPPLYSL--- 437  
 QY 368 GLIFPICDEAHVAVSVNRTPFOKKLMLVELNMNNSDQTEINMPLRATENPSTINGGF 427  
 Db 438 ---YMKDLTRAGATISGMV-AGAVVAVIWMISWIKPLA--TINAFPGMYEIIIP-----GF 484  
 QY 428 FDNVRLTFKGLTTMYLVVLLQFOISIPDKDSEGANNTIVDPVMSLD 480  
 Db 485 -----IVSVLTIYISKL-----TKKPD-----DVIENTLN 510

RESULT 11  
 US-09-949-016-10076  
 ; Sequence 10076, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIOR FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 10076  
 ; LENGTH: 1989  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-949-016-10076

Query Match 3.8%; Score 99.5; DB 4; Length 1989;  
 Best local similarity 19.8%; Pred. No. 1.1;  
 Matches 102; Conservative 80; Mismatches 177; Indels 155; Gaps 26;  
 QY 57 KYSIGLAERLDADYADPLDRKSSDS-TASNPPEKPSVYRNIDPIN-WFLRIIT----- 110  
 Db 1361 KYHCFMETSEIRIEDVNNKTECEKLMENGTIEIMKYNKINFDVNGAGYLLQVAT 1420  
 QY 111 --GVLPV-----RHGRPAKPEMNSAFYISVVFVLLACVGYANNRIHIVRSLSG 162  
 Db 1421 FKGMMDIMYAANDSRKDEQPKYEDNIYMYIYFVIFITP-----G 1460  
 QY 163 PFEBAVAVLVLVILPIMIPILMYEARKIAKLFNDWDFEVLVYQISGSLPLKLRQK 222  
 Db 1461 SF-----FTNLFIQVIL-----DNFNQKKKCGGDDIMTEBQK 1495  
 QY 223 AVYIA-----IVLPISLVSVITHTVMSDLINQVVPYCIIDNLTLMLG--- 267  
 Db 1496 KYVYAMKKGSKPKQKPIPRPLNKIQGIVFDFVTOQAFDIYIMMLICL--NMVTMMVETD 1553  
 QY 268 -----AMWFLICEAMSTIAHLAERFOKALKH-----IGPAMVADVVLMLRLSKL 314  
 Db 1554 TOSKQMENILYIMLVVVFIFTCVLMKF--ALHHYFTTG--WNIPDFVVALIST--- 1606  
 QY 315 TRDGNALCYTF--VFMSLYLFIITLISYGLMSQLSGF-GIKD-----IGLTTTALMNI 367  
 Db 1607 ---VGMPLADIIEKKYFSPITLFRVIRLARIGRLIRLIGAKGIRTLPLALMMSLPLAFNI 1663  
 QY 368 GLIFYI-----CDEAHVAVSVNRTPFO-----KKLMAVE 396  
 Db 1664 GLILFLVWIFISFGMSNFAVVKHAGIDIMFNPFETFGNSMILCFOITTSAGMDGLLPI 1723  
 QY 397 LNW---MNSDQTEINMPLRATENPSTINGGFVNRVTLFKGLTTMYLVVLLQFO 453

Db 1724 LNRPPDCLDKHKGSGF--KGDGNPSV---GIFVSVYIIISFLVYVNM-YIAIILE-N 1777  
 QY 454 ISIFTDKDGSEGANNTIVDP-----VMSLDND 482  
 Db 1778 FSVATE-----ESADPLSEDDFETFEYIEWEKFPD 1807

RESULT 12  
 US-09-538-092-564  
 ; Sequence 564, Application US/09538092  
 ; Patent No. 6753314  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glot, Loic  
 ; APPLICANT: Mansfield, Traci A.  
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
 ; FILE REFERENCE: 15966-542  
 ; CURRENT APPLICATION NUMBER: US/09/538,092  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: 60/127,352  
 ; PRIOR FILING DATE: 1999-04-01  
 ; PRIOR APPLICATION NUMBER: 60/178,965  
 ; PRIOR FILING DATE: 2000-02-01  
 ; NUMBER OF SEQ ID NOS: 1387  
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9  
 ; SEQ ID NO 564  
 ; LENGTH: 1900  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (0)..(0)  
 ; OTHER INFORMATION: Polypeptide Accession Number YLR305C  
 ; US-09-538-092-564

Query Match 3.7%; Score 99; DB 4; Length 1900;  
 Best local similarity 19.2%; Pred. No. 1.2;  
 Matches 90; Conservative 66; Mismatches 154; Indels 158; Gaps 20;

QY 70 YEAPPLDRKSSDSSTASNPPEKPSVYRNIDPINWFLRIITGVLPYRHGPAR----- 122  
 Db 727 YNSPPL-----ASFPNNKELS-----LEWNTILRRGSSNENIKQK 764  
 QY 123 -----AKEMNSAS---FIYSVVFVLLACVGYANNRIH--VRSLSGPFEE 166  
 Db 765 QQITEYNNINIVQRTTSSKIMFLAAVLELTRCAGDCSKLILFSDSILSGSIEK 824  
 QY 167 AVIAYFLVNLPIPIIPILMYEARKIAKLFNDWDFEVLVYQISGSLPLKLR-- 220  
 Db 825 CIA-----VLSVSMI-----RKYARLIQKNDALFNSKMTAQOLNMLLCISHREP 870  
 QY 221 --QKAVYIAVLPILSVSVITHTVMSDLINQVVPYCIIDNLTLMLGAMWFLICEAMS 278  
 Db 871 TLQDAAFACIEIFRISPSISLCHHLSL-----YTLIDMLTALFDS-----ILD 913  
 QY 279 ITAHLAERFOKALKHIGPAMV---ADYRVL-----MLR--LSGLTRDGNAL 322  
 Db 914 SEAKHPEPRYFKLHSHKTTIIVPSSSWRATTSLRLHKAKEVRIILNSNDT----- 969  
 QY 323 CYTFVMSLYLFIITLISYGLMSQLSGF-GIKDGLTTTALMNIIGLFIYICDEAHYA 380  
 Db 970 -----KILQSYISDGEYGRNLSVERGVSEFADMDAGLILPADKEISRLTYGPE----- 1019  
 QY 381 SVNVRTPFOKKLMLVELNMNNSDQTEINMPLRATENPSTINGGFDPVNRVTLFKGLLT 440  
 Db 1020 -----KPNITIS--GFTSLHSMRSKYLFD 1040  
 QY 441 TMVTVLVLLQFOISIPTD--KGDSEGANNTIVDPVMSDLNDMSLM 486  
 Db 1041 TAITSSPEDIRKQIGISTONIRKVLTLGNKTIITD-VTDFIDMATALL 1087

RESULT 13  
 US-09-489-039A-8212



Sequence 8212, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Brecot et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KIEBSTIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 8212  
LENGTH: 656  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8212

Query Match 3.7%; Score 98.5; DB 4; Length 656;  
Best Local Similarity 21.2%; Pred. No. 0.26;  
Matches 118; Conservative 67; Mismatches 206; Indels 165; Gaps 28;

QY 32 KDAVFLNKP-----LNSANQATLYGVRKTSIGLAERLDY--EAPPLRKKSSDST 84  
DB 75 KADAGLKNPQVIMVSEGGQFQVVGIVHVDVLAQVQAGLSDSEPAEERAA--- 131  
QY 85 ASNNPEKPSVFRNIDPINW-FLRIIGVLPV-----RHGPARAKF 125  
DB 132 -----KQSVLSQILIDITIGITFTPIGMAATGLKGLALAVTCMTPEOGYIKIWF 184  
QY 126 ENNSASFIYSVVFVLLACVGYVANNRI---HIVRSLSGFEBAVIAFLVNLIP-- 179  
DB 185 AASDLFF---FFPL---FLGYTAGKKEGKGNPFIEMVIGALTPLMIQAEASQAPGA 237  
QY 180 ----IMTPIILM--YEARIALFNDMDPEVLYIYQISGSHLPKLR---QKAVIATIL 230  
DB 238 AVEHPLGIVTFTINSSSVIPITLASW--VSCWLEKSNALLPSSKNKFFSPALCLAVV 295  
QY 231 P-I-L-S-V-L-S-V-I-T-H-V-M-S-D-I-N-Q-V-P-Y-C-I-L-D-N-I-L-T-A-M-L-G-A-M-W-F-----L-I-C-E 275  
DB 236 P-I-T-P-V-I-G-V-A-T-W-L-S-H-L-A-N-G-Y-O-F-I-Y-A-P-A-P-L-A-G-A-V-I-G-A-M-Q-V-C-Y-I-R-G-L-H-G-V-P-L-M-I-N 354  
QY 276 AMSITAHLLAERFQKALKHIGPAAVADYRVLMLRLSKLTRDTGNALCYTFVMSLYLFF 335  
DB 385 NMTVLGH-----DSWLPITLPAVIAQVAGVIGIFLA--TRDARQGVLAGSAF-SAGLFG 405  
QY 336 IITLSIYIGMSQSEGF-----GIKDIGLITLAL-----WNIQLFTICDEAHVASUV 384  
DB 406 ITPAIYGLTLPRLRPFFGCVAGA--IGGAIITAFSNSYASFGI----- 448  
QY 385 RTNFOKLLMVELNMNNSDAQTEINMFLRATENPSTINGGFPDVNRFLFKGLTTWT 444  
DB 449 -----PNIFFPAQMIIP-----GG---IDASVWGGLIGTGYA 477  
QY 445 Y-L-V-L-L-O-F-O-I-S-I-P-T-D-K-D-S-E-G-A-N-N-I-T-V-D-Y-W-D-S-I-D-N-M-S-----L-M 486  
DB 478 FVLACVLEFFAGLPR---GSAAPGAVTVAPASANDILAMSGSVIALAEQVDPSTFASGL 534  
QY 487 GASTLSTTTVGTTLPP 502  
DB 535 GKQVAIIPAVGOVIAP 550

RESULT 14  
US-08-671-525B-6  
Sequence 6, Application US/08671525B  
Patent No. 5703220  
GENERAL INFORMATION:  
APPLICANT: Yamada, Tadataka  
APPLICANT: Gantz, Ira  
TITLE OF INVENTION: Genes Encoding Melanocortin Receptors  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: US  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/671,525B  
FILING DATE: June 27, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Dean F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-000853DVB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810)641-1600  
TELEFAX: (810)641-0270  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-671-525B-6

Query Match 3.7%; Score 98; DB 1; Length 360;  
Best Local Similarity 21.3%; Pred. No. 0.12;  
Matches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17;

QY 32 KDAVFLNKP-----LNSANQATLYGVRKTSIGLAERLDYEAERPD 76  
DB 9 KEDFVFPVSSSFLRTLPQGLSALLTAMNASCCIPSVQPTLPNGSEHLQAPFS--- 64  
QY 77 RKSSDSTSANNPPEKPSVFRNIDPINWFLRIIGVLPVRHGPAPAKFENNSASFIYSV 136  
DB 65 --NQSSSACEQVFIKPEFL-SLGIVSLLENILVILAVVRGNLASHPMYF---FLCSL 117  
QY 137 VEFVLLACVGYVANNRIHIVRSLSGFEBAVIAFLVNLIPMIIPILMYEARKIATL 196  
DB 118 AVADMIVSVSNALFTIMAIIVASDYLTFFEDQIQH--MONIFDSMICISL-----VASI 169  
QY 197 FN-----DMDPEVLYIYQISGSHLPKLRQKAVIATILVILSVSVITHVMSDINQ 252  
DB 170 CNLLAIVADRYVITFYALRYSI-MTVRKALTLIVAIVCCGCVGF--IYSE--SK 223  
QY 253 VVPYCILDNLITAMLGAMWFLICEAMSTAHLLAERFQKALKHIGPAAVADYRVLMLRLS 312  
DB 224 NVIVCLITMFRMM-----LMGTLVHMFELPAKLHVKRIALPPADGVAPQGHSCMKGA 278  
QY 313 KLTRDTGNALCYTFVF---MSLYLFTIITLSIYIGMSQSEGFQIKDIGLTTIYALWNI 368  
DB 279 V-----TITILGCVFLFCMAPFLHLVLITCP-----TNPYCICYTAHFNTY 321  
QY 369 ILFYIC---DEAHYA--SVNVRITNFOKLLMVELNMN 401  
DB 322 LVLMCNVIDPLIYAFRSLELRNTFRE--ILCGCGNM 358

RESULT 15  
US-08-672-109B-6  
Sequence 6, Application US/08672109B  
Patent No. 5710265  
GENERAL INFORMATION:  
APPLICANT: Yamada, Tadataka  
APPLICANT: Gantz, Ira  
TITLE OF INVENTION: Genes Encoding Melanocortin Receptors  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: US  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/672,109B  
FILING DATE: June 27, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Dean F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-000853DVC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810)641-1600  
TELEFAX: (810)641-0270  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-672-109B-6

Query Match 3.7%; Score 98; DB 1; Length 360;  
Best Local Similarity 21.3%; Pred. No. 0.12;  
Matches 85; Conservative 64; Mismatches 172; Indels 78; Gaps 17;  
QY 32 KGDVAF-----LNKKPLNSANAQAYLYGVKYSIGLAERLDADYEAPPLD 76  
DB 9 EGDVFPVSSSSFLRTLEPQLSGALLTANASCLPSVQPTLPNGSEHLQAPFFS---- 64  
QY 77 RKSSDSSTASNNPEKPSVFYRNIDPINWFLRIIGVLPVHGPAPAKFENNSASFTYSV 136  
DB 65 --NQSSAPCEQYFIKPEIFL-SIGIVSLLENILVILAVVANGNLSPMYF----FLCSL 117  
QY 137 VFPEVLACYGYVANNRIHIVRSLSGPFEEAVIAYLPLVNIPLPIMILMYEARKAKL 196  
DB 118 AVADMIVSVSNALETIMIAIVHSDYLTFFEDPIQH--MDNIFDSMICISL-----VASI 169  
QY 197 FN-----DMDDPEVLYYOISGSHLPLKLRKAVYIAIVPLISVLSVYITHVMSDLNINQ 252  
DB 170 CNLLATVNDRYVITFYALRTHSI-MYRKALTLIVAIWCCGCGVVF--IVYSE---SK 223  
QY 253 VPEYCIIDNTIAMLGAMWFLICEAMSIYAHLLAERFQKALKHIGPAAVADRYVLMRLS 312  
DB 224 MVTVCILTFEFAMW-----LLMGTLYVHMFALRHVYRIALPAPDGVAPQHSCKMGA 278  
QY 313 KLTRDTGNALCYTFVF---MSLYLPIITLITSTIYGLMSQLESGFIDIGITITALNIG 368  
DB 279 V-----ITTIILGVPIFCWAPFFLVLITCP-----TNPYCICTAHFNTY 321  
QY 369 LLEFYIC---DEAHYA--SVVVRTNPOKCLMVEINMKN 401  
DB 322 LVLMCNVSVIDPLIYAFRSLELRNTFR--ILCGCNKN 358

Search completed: May 23, 2005, 19:26:45  
Job time : 46 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2005, 19:19:25 ; Search time 135 Seconds  
(without alignments)  
1268.651 Million cell updates/sec

Title: US-10-081-816-12  
Perfect score: 2645  
Sequence: 1 MRPSGEKVKVKGHGNSGHS.....TTTGTCTTTPPIKMKGRKG 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	2645	100.0	512	US-10-081-816-12	Sequence 12, Appl
2	390	14.7	364	US-10-081-816-24	Sequence 24, Appl
3	390	14.7	364	US-10-447-328-10	Sequence 10, Appl
4	169	6.4	33	US-10-081-816-67	Sequence 67, Appl
5	160	6.0	433	US-10-447-328-80	Sequence 80, Appl
6	135.5	5.1	369	US-10-081-816-29	Sequence 29, Appl
7	132.5	5.0	736	US-10-081-816-17	Sequence 17, Appl
8	130.5	4.9	372	US-10-081-816-30	Sequence 30, Appl
9	130.5	4.9	372	US-10-447-328-32	Sequence 32, Appl
10	129	4.9	436	US-10-081-816-11	Sequence 11, Appl
11	127	4.8	348	US-10-447-328-10	Sequence 10, Appl
12	122	4.6	410	US-10-081-816-1	Sequence 1, Appl
13	122	4.6	414	US-10-447-328-2	Sequence 2, Appl

14	120	4.5	450	US-10-081-816-57	Sequence 57, Appl
15	119	4.5	477	US-10-447-328-70	Sequence 70, Appl
16	118.5	4.5	2008	US-09-736-969A-2	Sequence 2, Appl
17	118.5	4.5	2008	US-09-736-969A-91	Sequence 91, Appl
18	118.5	4.5	2008	US-09-737-246-99	Sequence 99, Appl
19	118.5	4.5	2008	US-09-736-960-88	Sequence 88, Appl
20	118.5	4.5	2008	US-09-736-968A-105	Sequence 105, Appl
21	118.5	4.5	2008	US-09-978-244A-29	Sequence 29, Appl
22	115	4.3	367	US-10-081-816-13	Sequence 13, Appl
23	112.5	4.3	365	US-10-081-816-56	Sequence 56, Appl
24	112.5	4.3	408	US-10-447-328-50	Sequence 50, Appl
25	110.5	4.2	408	US-10-081-816-37	Sequence 37, Appl
26	110.5	4.2	2073	US-09-978-244A-10	Sequence 10, Appl
27	110	4.2	498	US-10-081-816-15	Sequence 15, Appl
28	108	4.1	381	US-10-081-816-31	Sequence 31, Appl
29	108	4.1	381	US-10-447-328-34	Sequence 34, Appl
30	107.5	4.1	346	US-10-282-122A-77576	Sequence 77576, A
31	107	4.0	404	US-10-081-816-52	Sequence 52, Appl
32	106.5	4.0	390	US-10-081-816-39	Sequence 39, Appl
33	106.5	4.0	390	US-10-447-328-58	Sequence 58, Appl
34	106.5	4.0	702	US-10-474-776-384	Sequence 384, App
35	106.5	4.0	702	US-10-472-928-4070	Sequence 4070, Ap
36	106.5	4.0	816	US-09-252-088-5	Sequence 5, Appl1
37	106.5	4.0	816	US-10-340-792-5	Sequence 5, Appl1
38	104	3.9	363	US-10-601-309-74	Sequence 74, Appl1
39	103.5	3.9	960	US-10-214-390-4	Sequence 4, Appl1
40	103	3.9	287	US-10-081-816-22	Sequence 22, Appl1
41	103	3.9	1011	US-10-320-797-3324	Sequence 3324, Ap
42	102.5	3.9	373	US-10-081-816-4	Sequence 4, Appl1
43	102.5	3.9	373	US-10-447-328-8	Sequence 8, Appl1
44	102	3.9	320	US-09-510-332-160	Sequence 160, App
45	102	3.9	320	US-10-962-365-160	Sequence 160, App

## ALIGNMENTS

RESULT 1  
US-10-081-816-12  
; Sequence 12, Application US/10081816  
; Publication No. US20030045472A1  
; GENERAL INFORMATION:  
; APPLICANT: Axel, Richard  
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto  
; FILE REFERENCE: 0575/64019-A/jpw/ADM  
; CURRENT APPLICATION NUMBER: US/10/081,816  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/271,319  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-081-816-12

Query Match 100.0%; Score 2645; DB 14; Length 512;  
Best Local Similarity 100.0%; Pred. No. 2,1e-252;  
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPSGEKVKVKGHGNSGHSLSGNANYRRKKGDVFLNAKPLNSANAQATLYGRTKSI 60  
DB 1 MRPSGEKVKVKGHGNSGHSLSGNANYRRKKGDVFLNAKPLNSANAQATLYGRTKSI 60  
QY 61 GLARLDVYAPPLDRKSSDSTASNNPERPSGFYNDIPIMFLRIIGVLPVIRGCP 120  
DB 61 GLAEPLDVYAPPLDRKSSDSTASNNPERPSGFYNDIPIMFLRIIGVLPVIRGCP 120  
QY 121 ARAKENMSASFISVVFVLACVGVANRRITHIVSLSGPEEAVIAYFLVNIPI 180

```

Db 121 ABAKEMNSAFIYSVVFVLLACVGVYVANNRHHVRSLSGFEBAVAVLVLVILPI 180
Qy 181 MIPIIMYARKIAKLFNDMDPEVLXYOISGSLPLKROKAVYAIYLPILSVSVYI 240
Db 181 MIPIIMYARKIAKLFNDMDPEVLXYOISGSLPLKROKAVYAIYLPILSVSVYI 240
Qy 241 THVMSDLINOVVPYCIIDNTLAMLGAMWFLICEAMSTIHLAERFOKALKHIGPAM 300
Db 241 THVMSDLINOVVPYCIIDNTLAMLGAMWFLICEAMSTIHLAERFOKALKHIGPAM 300
Qy 301 VADRYVLMRLSKTRDTGNALCYTFVMSLYIFITTSYGLMSQSLSEGGIKDIGIT 360
Db 301 VADRYVLMRLSKTRDTGNALCYTFVMSLYIFITTSYGLMSQSLSEGGIKDIGIT 360
Qy 361 ITALMNIGLIFYICDEAHYASVVRNTPFOKKLMLVNLNMSDAOTEINMPLRATENPS 420
Db 361 ITALMNIGLIFYICDEAHYASVVRNTPFOKKLMLVNLNMSDAOTEINMPLRATENPS 420
Qy 421 TINCQGFDPVNRTEFKGLITTVWTVLVLLQFOISIPTRKDSBEGANNITVVDVWDSID 480
Db 421 TINCQGFDPVNRTEFKGLITTVWTVLVLLQFOISIPTRKDSBEGANNITVVDVWDSID 480
Qy 481 NDMSLMGASTLSTTTVGTTLPPIMKLKGRKG 512
Db 481 NDMSLMGASTLSTTTVGTTLPPIMKLKGRKG 512

```

## RESULT 2

```

US-10-081-816-24
; Sequence 24, Application US/10081816
; Publication No. US20030045472A1
; GENERAL INFORMATION:
; APPLICANT: Scott, Kristin
; APPLICANT: Axel, Richard
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 05/75/64019-A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-081-816-24

```

```

Query Match 14.7%; Score 390; DB 14; Length 364;
Best Local Similarity 27.3%; Pred. No. 1.4e-29;
Matches 101; Conservative 77; Mismatches 152; Indels 40; Gaps 9;

```

```

Qy 110 IGVLPVIRHGP---ARAKFEMNSAFIYSVVFVLLACVGVYVANNRHHVRSLSGPF 164
Db 110 IGVLPVIRHGP---ARAKFEMNSAFIYSVVFVLLACVGVYVANNRHHVRSLSGPF 164
Qy 165 BEAVIAYLVLVNLIPILMIPIL-WYEARKIAKLFNDMDPEVLXYOISGSLPLKLRQKA 223
Db 165 BEAVIAYLVLVNLIPILMIPIL-WYEARKIAKLFNDMDPEVLXYOISGSLPLKLRQKA 223
Qy 224 VYIAIVLPILSVLS---VITHTVMSDLINOVVPYCIIDN-----LTAMLGAM 269
Db 224 VYIAIVLPILSVLS---VITHTVMSDLINOVVPYCIIDN-----LTAMLGAM 269
Qy 270 ---WFLICAMSTIHLAERFOKALKHIGPAMVADVVLMLRSLKTRDGNALCYTF 326
Db 270 ---WFLICAMSTIHLAERFOKALKHIGPAMVADVVLMLRSLKTRDGNALCYTF 326
Qy 326 GMYCLVIFPTTIIATYGSISEIIDHGATYKEVGLFVIYVYCGGLYIICNEAHYASRKVG 284
Db 326 GMYCLVIFPTTIIATYGSISEIIDHGATYKEVGLFVIYVYCGGLYIICNEAHYASRKVG 284

```

```

Qy 386 TNFOKKLMLVNLNMSDAOTEINMPLRATENPSTINCQGFDPVNRTEFKGLITTVWTV 445
Db 386 TNFOKKLMLVNLNMSDAOTEINMPLRATENPSTINCQGFDPVNRTEFKGLITTVWTV 445
Qy 446 LVVLLQFOIS 455
Db 446 LVVLLQFOIS 455
Qy 345 LVVLLQFOKIT 354
Db 345 LVVLLQFOKIT 354

```

## RESULT 3

```

US-10-447-328-10
; Sequence 10, Application US/10447328
; Publication No. US20040003419A1
; GENERAL INFORMATION:
; APPLICANT: Carlgren, John R.
; APPLICANT: Clyne, Peter J.
; APPLICANT: Warr, Coral G.
; APPLICANT: Yale University
; TITLE OF INVENTION: No. US20040003419A1 Taste Receptors in Drosophila
; FILE REFERENCE: 44574-5072
; CURRENT APPLICATION NUMBER: US/10/447,328
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/09/593,519
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,668
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 60/181,704
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-447-328-10

```

```

Query Match 14.7%; Score 390; DB 15; Length 364;
Best Local Similarity 27.3%; Pred. No. 1.4e-29;
Matches 101; Conservative 77; Mismatches 152; Indels 40; Gaps 9;

```

```

Qy 110 IGVLPVIRHGP---ARAKFEMNSAFIYSVVFVLLACVGVYVANNRHHVRSLSGPF 164
Db 110 IGVLPVIRHGP---ARAKFEMNSAFIYSVVFVLLACVGVYVANNRHHVRSLSGPF 164
Qy 165 BEAVIAYLVLVNLIPILMIPIL-WYEARKIAKLFNDMDPEVLXYOISGSLPLKLRQKA 223
Db 165 BEAVIAYLVLVNLIPILMIPIL-WYEARKIAKLFNDMDPEVLXYOISGSLPLKLRQKA 223
Qy 224 VYIAIVLPILSVLS---VITHTVMSDLINOVVPYCIIDN-----LTAMLGAM 269
Db 224 VYIAIVLPILSVLS---VITHTVMSDLINOVVPYCIIDN-----LTAMLGAM 269
Qy 270 ---WFLICAMSTIHLAERFOKALKHIGPAMVADVVLMLRSLKTRDGNALCYTF 326
Db 270 ---WFLICAMSTIHLAERFOKALKHIGPAMVADVVLMLRSLKTRDGNALCYTF 326
Qy 326 GMYCLVIFPTTIIATYGSISEIIDHGATYKEVGLFVIYVYCGGLYIICNEAHYASRKVG 284
Db 326 GMYCLVIFPTTIIATYGSISEIIDHGATYKEVGLFVIYVYCGGLYIICNEAHYASRKVG 284
Qy 345 LVVLLQFOKIT 354
Db 345 LVVLLQFOKIT 354

```

## RESULT 4

```

US-10-081-816-67
; Sequence 67, Application US/10081816
; Publication No. US20030045472A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Axel, Richard
; APPLICANT: Scott, Kristin
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recept
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 0575/64019-A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-081-816-67

Query Match      6.4%; Score 169; DB 14; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      420 STINGGFPDYNRTLFKGLTTWTYVVLQF 452
DB      1 STINGGFPDYNRTLFKGLTTWTYVVLQF 33

RESULT 5
US-10-447-328-80
; Sequence 80, Application US/10447328
; Publication No. US20040003419A1
; GENERAL INFORMATION:
; APPLICANT: Carlson, John R.
; APPLICANT: Clyne, Peter J.
; APPLICANT: Mair, Coral G.
; APPLICANT: Yale University
; TITLE OF INVENTION: No. US20040003419A1 Taste Receptors in Drosophila
; FILE REFERENCE: 44574-5072
; CURRENT APPLICATION NUMBER: US/10/447,328
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/09/593,519
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,668
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 60/181,704
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-447-328-80

Query Match      6.0%; Score 160; DB 15; Length 433;
Best Local Similarity 20.9%; Pred. No. 1e-06;
Matches 91; Conservative 72; Mismatches 139; Indels 134; Gaps 18;

QY      102 PIMFLAIIIGLPIVRHGPAPAKFEMNSASFIYVFFVLLACVGYVANNR-IHIVRS- 159
DB      27 PIGFTFTL-----HRRRGIVITLGYACYLISISLWVTECYANVALQKHKTRAE 79

QY      160 ----LSGPF--EAVIAYFLVNIILPIMIPILWYBARKIAKLFDNDWDFEV-LYYQISG- 212
DB      80 DSSKVMBEVARSMMVAMFVWNOINILL-----NFRRLARIYDDIDLLEIDLNNASGF 133

QY      213 ----HSLPLKLR-QKAVYIAVL-----PILSVSVITHTVMSDLINOV 253
DB      134 VGRHWRFRFRRLISVGLMTVLVGLTPRPTVALGFIYLMTKVLTETIILMLQI-KC 192

QY      254 VPYICILNLITAMLGAMWFLICEAMSTAHLABRFQKALKHIGPAWADVRLMLRLSK 313
DB      193 TEYCV-----FVLLIYELIRGRHIL-QQISVLE-----GN 223
```

```

QY      314 LTRDTGNALCYTFVMSLYLPIITLISYGLMSQLSBEGIKDIGLITLALNNIGLFLY 373
DB      224 QSRDSVOELC-----VALKRNQDLAGRIGLVNVEVLYF---TSLTLFLVNEHLITLOI 275

QY      374 CDBAHYASVN-----VRTNFOKLMVBLNMWNSDAQETINNELRA--TEWN 418
DB      276 VNALIKSVNPNCCQYSKLVFEKKNFTYKQVIFIIIGVGTCLLSINIFLSCLVSEFC 335

QY      419 PSTIN-----CGGPDVNRITLFGKL 438
DB      336 IQYNSISRVLHQMYCISAEDYLILKMGIREYSLQMEHLKLIPTCGGLDINLKFEQGM 395

QY      439 LTTMTYVVLVLOFOI 454
DB      396 VVTLFGYIILVQPKI 411

RESULT 6
US-10-081-816-29
; Sequence 29, Application US/10081816
; Publication No. US20030045472A1
; GENERAL INFORMATION:
; APPLICANT: Axel, Richard
; APPLICANT: Scott, Kristin
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor
; FILE REFERENCE: 0575/64019-A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-081-816-29

Query Match      5.1%; Score 135.5; DB 14; Length 369;
Best Local Similarity 20.5%; Pred. No. 0.00021;
Matches 87; Conservative 65; Mismatches 126; Indels 147; Gaps 20;

QY      106 FLRIIGVLPVIRHGPAPAKFEMNSASFIYVFFVLLACVGYVANNRIHVRLSGPFE 165
DB      11 YFALGLVLP--WSBSCAQSKFVQK-----VYSAILLITL-----NAVHGISIYEP-Q 54

QY      166 EAVIAYFLVNIILPIMIPILWYBARKIAKLFDNDWDF-----EVLYY-----QI 210
DB      55 SAEFLSLMNVAVIYFVARIVCATVI-----ILQVWVHIDYRFGCEMKTGLGRLOCEL 108

QY      211 SGHSLPKLRQKAVYIAI-----VLPIIL-----SVLSVIT- 241
DB      109 KIHVGRILKQSYAKIILGLGFLVTVPISYVVALSGSLVYFMSLLILIRMQFVLVL 168

QY      242 -----HTWMSDLINOV-----DYCLIDNLTAMLGAMWFLICEAMS-ITNHLAER 287
DB      169 NVELLGRHVSILGIRLQNVLECHLMGANCTLDGANRLCSLEFLALAKQSMQAHYLFTH 228

QY      288 FQKALKHIGPAWADVRLMLRLSKLTPRTGNALCYTFVMSLYLFFITLSYGLMSQ 347
DB      229 FNDLP-----GMSILGTV-----VLSDSIVNIYTWQOV 258

QY      348 LSEGGIKDIGLT-----ITALMNIGLFLYICD-----AHYASVNRVN 387
DB      259 LVEYVEYKLYATATSVVPSPFN-LVFCRCGECQQRQSVLYIGSYLRNLISCHPSIGRETS 317

QY      388 FQKGLMVELAMWNSDAQETINNELFRLATEWNPSTINGGFPDVRITLFGKLLTTMTYV 447
DB      318 Y-KDLM-----EPILQVEQVLAINEGFMSTONSILMSITAAKVYILI 361

QY      448 VLLQF 452
```

Db 362 VLMQF 366

RESULT 7  
US-10-081-816-17  
; Sequence 17, Application US/10081816  
; Publication No. US20030045472A1  
; GENERAL INFORMATION:  
; APPLICANT: Axel, Richard  
; APPLICANT: Scott, Kristin  
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto  
; FILE REFERENCE: 0575/64019-A/JPM/ADM  
; CURRENT APPLICATION NUMBER: US/10/081, 816  
; CURRENT FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/271,319  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-081-816-17

Query Match 5.0%; Score 132.5; DB 14; Length 736;  
Best Local Similarity 21.0%; Pred. No. 0.0011; Matches 77; Conservative 66; Mismatches 134; Indels 89; Gaps 16;

QY 102 PINMFLRIIGVLPYVHGPAPAKFEMNSASFTYSVVFVLLACVYGVYANNR-IHIVRS- 159  
DB 427 PIGQFTRRL-----HKRRRGIVILGYACYLISLIMVYEGYANIVLQDKHKEHAB 479  
QY 160 ---LSGPEEAVIAYLVFNVLPIPTIITILYTPARKIAKLFNDMDPEV-LYQISG-- 212  
DB 480 DSSKTMGNQKVLVAMFVWNOILNL-----NFRRLRIYDIDADEIDLNASSGFEV 533  
QY 213 ---HSLPLKLR-QKAVYIAIVL-----PISVLSVVTHTYMTSDINNOV 254  
DB 534 GGRHMMRRFRRLALSLVGLWITVLVGLTPFTLVALCPYLAHTNKKVLTETIILMLQI-KCT 592  
QY 255 PYCILDNLTMAGMWFLICEAMSTYAHLLAEKFOALKHIGPAAVADYRVLMRLSKL 314  
DB 593 EXCV-----FVLLIYELILRGRIHL-QQISVEL-----GNQ 623  
QY 315 TRDTGNALCYTVEVMSLYLFTITLSTYGLMSQLSGFGIKDIGITTLTALNIGLLFYIC 374  
DB 624 SBDSSVQELC-----VALKENQILAGRIWGLVNEVSLYF---TSLTLFLVYNELTILQIV 675  
QY 375 DEAHYASVNVV--TNPQKLLMVELMNMNSDAQTEINMFLRATENMPSTINGCGFPDVR 432  
DB 676 MWALLKSVNPNECCQYTEDYILIKMGLRYSIQMEHLIKI-----FTCGGLFDIML 726  
QY 433 TLFKGL 438  
DB 727 KFFGGV 732

RESULT 8  
US-10-081-816-30  
; Sequence 30, Application US/10081816  
; Publication No. US20030045472A1  
; GENERAL INFORMATION:  
; APPLICANT: Axel, Richard  
; APPLICANT: Scott, Kristin  
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto  
; FILE REFERENCE: 0575/64019-A/JPM/ADM  
; CURRENT APPLICATION NUMBER: US/10/081, 816  
; CURRENT FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/271,319  
; PRIOR FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-081-816-30

Query Match 4.9%; Score 130.5; DB 14; Length 372;  
Best Local Similarity 21.7%; Pred. No. 0.00067; Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;

QY 129 SASFTYSVVFVLLACVYGVYANNRIHI-VRSLSG-PPEEAVI-----AYLFL--- 174  
DB 39 SSTAIQIVVGVFMALGALASLYMETSKQNTGFNNAVILTTSVQQLANMLRSG 98  
QY 175 ---VNILPIM--IIPILYTPARKIAKLFNDMDPEVLYIQISGSLPLKQKAVYIAI 228  
DB 99 QKSQVNLQRLSQVVELLQFEFYAVPQ-----FRMLY-----RI 132  
QY 229 VPIILSVLSVVTHTYMTSDINNOV-----PYCILDNLTMAGMWFLICEAMSI 279  
DB 133 WLTVCLITGAMVTHGIMNLTTMQISRVLTLLIGFYRCVLAN-----FQFTCYGMV 185  
QY 280 TAILAEERFOKALKHI-----GPAAVADY-RVLMRLSKLTRDTGNALCYTVEF- 328  
DB 186 ILKDLQVQKQLEHLVSTTISMGVAGCLATHDEIILLGRELIAVYGVILFLFIYQ 245  
QY 329 --MSLYLFFITLSTYGLMSQLSGFGIKDIGITTLTALNIGLLFYICDEAHYASVNVRT 386  
DB 246 VMOCLILFIYSNL-----EGFHSN-DLVLIFCWLAHMLPYLLIPLVNDIHNOA 294  
QY 387 NFQKLLMVELMNMNSDAQTEINMFLRATENMPSTINGCGFPDVR-TLFKGLITMTVY 445  
DB 295 NKTAKML-TKVPRGTGLDRMIKFLKRLKQKPIITAYGFAIDKSTILFK-LFTRIFTY 352  
QY 446 LVVLQFPQ 453  
DB 353 WVLVQFK 360

RESULT 9  
US-10-447-328-32  
; Sequence 32, Application US/10447328  
; Publication No. US20040003419A1  
; GENERAL INFORMATION:  
; APPLICANT: Carlson, John R.  
; APPLICANT: Clyne, Peter J.  
; APPLICANT: Warr, Coral G.  
; APPLICANT: Yale University  
; TITLE OF INVENTION: No. US20040003419A1el Taste Receptors in Drosophila  
; FILE REFERENCE: 44574-5072  
; CURRENT APPLICATION NUMBER: US/10/447,328  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: US/09/593,519  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/138,668  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 60/181,704  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-447-328-32

Query Match 4.9%; Score 130.5; DB 15; Length 372;  
Best Local Similarity 21.7%; Pred. No. 0.00067; Matches 80; Conservative 64; Mismatches 135; Indels 89; Gaps 17;

QY 129 SASFTYSVVFVLLACVYGVYANNRIHI-VRSLSG-PPEEAVI-----AYLFL--- 174



```

Db      39 SSTAIGIVVGVPMALLGLAELSLYMETKSGTGTEDNAVILTTSTVOLLANMLRSQ 98
Qy      175 ----VNILPIM--IIPIMYEAARKIACLFENDWDFEVLYYOISGHSPLKLRKAVYIAI 228
Db      99 QKSOVNLQSLQSVLLOFEPAVAPQ-----FRLLY-----R1 132
Qy      229 VLPILSVLAVVITHTVMSDINQV-----VPCIIDNLTAMLGAMWFLICEAMSI 279
Db      133 WLLVLCIYGAMVTHFGINMLTWMQISRVLTIGFVACVLAN-----PQFCYGTGMV 185
Qy      280 TAILLAERFOKALKH-----GPAWVADY-RVLMRLSLKLTGDNALCYTFVF- 328
Db      186 ILKKLQOVQKQLEHVTSTTISMAVAGGLRTIDEILLQORELAVYGGVILFLFIYQ 245
Qy      329 --MSLYLFIITLTIYGLMSQSLSEFGIKDIGLITLTMNIGLFYICDEAHYAVAVRT 386
Db      246 VMOCILIFYSNL-----EGFHSSN-DLVLCICMLAPMLFYILPLVVDINHQA 294
Qy      387 NFOKRLMVELMNMNSDAQEINMFLPATENPSTINGCFPDVNR-TLFRGLITTVTY 445
Db      295 NKTAKML-TVPRGTGLDMIRKFLKNLRKQKILTAYGFPALDKSTLFR-LFTALFTY 352
Qy      446 LVYLLQRF 453
Db      353 MWLVQPK 360

```

## RESULT 10

US-10-081-816-11

```

; Sequence 11, Application US/10081816
; Publication No. US20030045472A1
; GENERAL INFORMATION:
; APPLICANT: Axel, Richard
; APPLICANT: Scott, Kristin
; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recepto
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 0575/64019-A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/10/081,816
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/271,319
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-081-816-11

```

Query Match 4.9%; Score 129; DB 14; Length 436;  
 Best Local Similarity 21.2%; Pred. No. 0.0012;  
 Matches 94; Conservative 70; Mismatches 172; Indels 108; Gaps 22;

```

Qy      85 ASNNPERK-SVFRNNDPIWMLRIIGVLP--LVHGRPARAKEMNSASFIYSVVFV 140
Db      30 AOGGIEFEQDLTFYGAIRPYLCVAQFGEIMPLSNIRSDPOVRFKXNSIGLAVTGFL 89
Qy      141 L--LACVGVYANNRIRHIVRSLSGPFBEAVIAYFLVNLPIIMIIPIIMYE--ARKIA 194
Db      90 LGCKMTLVG--AN-----ILFTBGLMAKNIYGLVFLI-----VGMVWMLNFGPARSWS 136
Qy      195 KLFNDMDFEVLY--YQISGHSPLKLRKQKAVYIAIVLPILSV----- 235
Db      137 HMLPMSVDILMLFPYKRGKRS---LRKVNVALSVVLAAGDMELTYASGYSYS 192
Qy      236 LSVVITHTVMSD-----NINQVPCILNDITAM-LGAW-----WELICE 275
Db      193 WHIQCCHNHSRIITFGVLEKEFSDIMFMP---NIFSMCYGFWLGAFTFLMFMFI 248
Qy      276 AMSITAHLLARFOKALKHIG-----PAWVADYRVLMRLSLGLTRDTGNALCYTFVF 328
Db      249 PIWVTSIGLQRFQGFARVAGLBSRHVPEALMTDIRDHIRCELAA-----SLVEAS 301

```

```

Qy      329 MSYLFIITLTIYGLMSQSLSEFGIKDIGLITLTM-----NIGLFYICDEAHYA 380
Db      302 MSNIVFVSCANNVYVINCQALALFTYLRHPINIVYVMSILFLLARTSLVFTASKHDA 361
Qy      381 SVAVRTNFOKRLMVELMNMNSDAQEINMFLPATENPSTINGCF--FDVNRTLFKG 437
Db      362 SL-----LRLRLYLVPISDQWTOEVORF--ADQLISEFVGLSGYRLFCLTRKSLFG 410
Qy      438 LTTMTYTVLLVLLQFQISIFPDKG 461
Db      411 MLATLVTYEIMLP--QIDAKSHKG 432

```

## RESULT 11

US-10-447-328-30

```

; Sequence 30, Application US/10447328
; Publication No. US20040003419A1
; GENERAL INFORMATION:
; APPLICANT: Carlson, John R.
; APPLICANT: Clyne, Peter J.
; APPLICANT: Wait, Coral G.
; APPLICANT: Yale University
; TITLE OF INVENTION: No. US20040003419A1el Taste Receptors in Drosophila
; FILE REFERENCE: 44574-5072
; CURRENT APPLICATION NUMBER: US/10/447,328
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/09/593,519
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,668
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 60/161,704
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-447-328-30

```

Query Match 4.8%; Score 127; DB 15; Length 348;  
 Best Local Similarity 19.8%; Pred. No. 0.0013;  
 Matches 81; Conservative 62; Mismatches 129; Indels 138; Gaps 17;

```

Qy      106 FLRIIGVLPVIRHGRPARAKEMNSASFIYSVVFVLIACY----- 145
Db      11 VFALGLVLP--WSECAQSKFVQK---VSGALILITLNAVHGISIVFPQSAELFLSIMGN 65
Qy      146 -VGTV-----ANNRHHIVRSLSGPFBEAV-IAYFLVNLPIIMIIPIIMYE 189
Db      66 GIGVTRIAQGTGYGLRLQCELRKIHVERLKMOSYAKLILAGIGLVTVLPSTIVALL 121
Qy      122 -----SSLSIYFMSLSLST-LIIRMQVVLVLANVELG-----HHYSLSGIR 162
Db      250 INQVY-----PYCIIDNLTAMLGAMWFLICEAMS-ITAHLLARFOKALKHIPAAWVA 302
Qy      163 LQNVLECHLMGANCCTLDGNANRLCSLFEFLALAKOSHQOLHFLTFHFNDLF----- 212
Db      303 DYRVLMRLSLGLTRDTGNALCYTFVPMISVLYFIITLTIYGLMSQSLSEFGIKDIGLT- 360
Qy      213 -----GMSITGTV---VLFSDSTVNIWTQOVLVEVIEYKLYATFS 252
Db      361 --ITALNIGLFIYICD-----AHYASVAVRTNFOKRLMVELMNMNS 402
Qy      253 VFVPSFRTN-LVFRGCEFPQORQSVLIGSYARNLSCHPSIGRETSY-KDILM----- 302
Db      403 DAQTEINMFLPATENPSTINGCFPDVNRTLFPGLLITTVTYIVYLLQF 452
Qy      303 -----EFILQVEQNVYLAIVABGFMSDMSILMSIILAAKYTVLIVMQF 345

```

RESULT 12  
 US-10-081-816-1  
 ; Sequence 1, Application US/10081816  
 ; Publication No. US20030045472A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Axel, Richard  
 ; APPLICANT: Scott, Kristin  
 ; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Recept  
 ; TITLE OF INVENTION: Uses Thereof  
 ; FILE REFERENCE: 0575/64019-A/JPM/ADM  
 ; CURRENT APPLICATION NUMBER: US/10/081,816  
 ; PRIOR FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: 60/271,319  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 116  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 410  
 ; TYPE: PRF  
 ; ORGANISM: Drosophila melanogaster  
 US-10-081-816-1

Query Match 4.6%; Score 122; DB 14; Length 410;  
 Best Local Similarity 22.2%; Pred. No. 0.0053;  
 Matches 70; Conservative 53; Mismatches 104; Indels 88; Gaps 18;

QY 185 ILM-YEARKI-----AKLFMDWDFEVLYYQISGSLPL---KLRKAVYIAIYLPILSTL 236  
 DB 149 ILMGAVSQLLILGALSLRGRFPF--YWIS-YLLPLVCGRLRYQIFNAQL-VRORL 204  
 QY 237 SVVITHTVMSDININOVPEYCIIDNLTAMLGAWMFLICEAMSTAHILAEPOKALKHIG 296  
 DB 205 DVLL--VALQQLQHLQKGP-----AVDTVLEQDLEBAADRIL 242  
 QY 297 PAAWADYVWLRLSKLTRDTGNALCYTFVMSLYL-----FFITLSIYGLMSQLSSEG 351  
 DB 243 AVRLLV--YORVVALVALNLR-----CYG---LSMLQVGNDFLATTSNCYWMFLNFRQS 291  
 QY 352 ----FGIKDIGLITLWM-----NIGLLFYICDEA-----HYAVNVRNTPOK 390  
 DB 292 AASPFI--LQIVASGVWSAPHLGNTVLVSLCDRTAQCASRLALCLHQVSDLRNESHN 349  
 QY 391 KLIMVELMNMNSDAQTEINMFLRATENPSTINCSPFPDVRNRLFGLLTTWVTVLVLL 450  
 DB 350 ALI-----TQFSLQLLHQRLH---FSAAGFVNDCTLTATYIGATTYYIIILI 394  
 QY 451 QFOISIFPTDKGDSSEG 465  
 DB 395 QFHMSESTIGSDSNG 409

RESULT 13  
 US-10-447-328-2  
 ; Sequence 2, Application US/10447328  
 ; Publication No. US20040003419A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Carlson, John R.  
 ; APPLICANT: Warr, Coral G.  
 ; APPLICANT: Clayne, Peter J.  
 ; APPLICANT: Yale University  
 ; TITLE OF INVENTION: No. US20040003419A1e1 Taste Receptors in Drosophila  
 ; FILE REFERENCE: 44574-5072  
 ; CURRENT APPLICATION NUMBER: US/10/447,328  
 ; CURRENT FILING DATE: 2003-05-29  
 ; PRIOR APPLICATION NUMBER: US/09/593,519  
 ; PRIOR FILING DATE: 2000-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/138,668  
 ; PRIOR FILING DATE: 1999-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/181,704  
 ; PRIOR FILING DATE: 2000-02-10  
 ; NUMBER OF SEQ ID NOS: 94  
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2  
 ; LENGTH: 414  
 ; TYPE: PRF  
 ; ORGANISM: Drosophila melanogaster  
 US-10-447-328-2

Query Match 4.6%; Score 122; DB 15; Length 414;  
 Best Local Similarity 22.2%; Pred. No. 0.0054;  
 Matches 70; Conservative 53; Mismatches 104; Indels 88; Gaps 18;

QY 185 ILM-YEARKI-----AKLFMDWDFEVLYYQISGSLPL---KLRKAVYIAIYLPILSTL 236  
 DB 153 ILMGAVSQLLILGALSLRGRFPF--YWIS-YLLPLVCGRLRYQIFNAQL-VRORL 208  
 QY 237 SVVITHTVMSDININOVPEYCIIDNLTAMLGAWMFLICEAMSTAHILAEPOKALKHIG 296  
 DB 209 DVLL--VALQQLQHLQKGP-----AVDTVLEQDLEBAADRIL 246  
 QY 297 PAAWADYVWLRLSKLTRDTGNALCYTFVMSLYL-----FFITLSIYGLMSQLSSEG 351  
 DB 247 AVRLLV--YORVVALVALNLR-----CYG---LSMLQVGNDFLATTSNCYWMFLNFRQS 295  
 QY 352 ----FGIKDIGLITLWM-----NIGLLFYICDEA-----HYAVNVRNTPOK 390  
 DB 296 AASPFI--LQIVASGVWSAPHLGNTVLVSLCDRTAQCASRLALCLHQVSDLRNESHN 353  
 QY 391 KLIMVELMNMNSDAQTEINMFLRATENPSTINCSPFPDVRNRLFGLLTTWVTVLVLL 450  
 DB 354 ALI-----TQFSLQLLHQRLH---FSAAGFVNDCTLTATYIGATTYYIIILI 398  
 QY 451 QFOISIFPTDKGDSSEG 465  
 DB 399 QFHMSESTIGSDSNG 413

RESULT 14  
 US-10-081-816-57  
 ; Sequence 57, Application US/10081816  
 ; Publication No. US20030045472A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Axel, Richard  
 ; APPLICANT: Scott, Kristin  
 ; TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptor  
 ; TITLE OF INVENTION: Uses Thereof  
 ; FILE REFERENCE: 0575/64019-A/JPM/ADM  
 ; CURRENT APPLICATION NUMBER: US/10/081,816  
 ; CURRENT FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: 60/271,319  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 116  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 57  
 ; LENGTH: 450  
 ; TYPE: PRF  
 ; ORGANISM: Drosophila melanogaster  
 US-10-081-816-57

Query Match 4.5%; Score 120; DB 14; Length 450;  
 Best Local Similarity 19.0%; Pred. No. 0.0096;  
 Matches 91; Conservative 76; Mismatches 149; Indels 162; Gaps 26;

QY 71 EAPPLDRKKSSTASNN-----PEF-KPSVFRNIDPIMWFLRIIGVLPV--RHG 119  
 DB 28 ETPP--PKVEDSNLEFNVLASBKLPNTYNLDLFRAVPPFVFLAQCAVIMPLVGIRESN 85  
 QY 120 PARAKFEMNSASFYSVVFV-----LLACY-----VGVANRRHIYVSLSGPFEAVY 169  
 DB 86 PRVRVRAKYSIPMFVYLLIFMIAISILFLSMFTHLKGITAKNFVGLV--FQCVLS 140  
 QY 170 AYLFVNIPLPIMIPILMYEARKIAKLFPMDWDFEVL---YQISGSLP----- 216  
 DB 141 AYVVFIRL-----AKKPAVAVVIMVTRTEIPFKPPEIPKRLISRRVQLAALA 188

```
Qy 217 ---LKKRQKAVY--IAIVLPILSVLVITHTVMSDLN-----INQVVPY-----C 257
Db 189 IIGSLGEGHLYQVSALSYTRRIQMCANITTVSPFNMYNOTNDYVQOLPSPITAVL 248
Qy 258 IILNLTAMLGAMWFLICEAMSTIAHLAERFOKALKHIGPAAWADYRVLMRLSKLTRD 317
Db 249 ILATCTIV---NNYMDLFIMMISKG--LSYREQ-----ITTRIRKLEHE 288
Qy 318 TGNALCYTFVPMISLYLPITTLISYGLMSQLSGFGIKDIGTLTALMN--IGLLFYICDE 376
Db 289 ---EVCS-----VFQIREHYVKCELE---FVDSAMSSLIILSCVNNLYFVC-- 332
Qy 377 AHYASVVRNTRFOCKLMLVELNW-----MNSDAQREIN 409
Db 333 --YQLNIV---FNK-----LRMPINITYFWYSLLYLIGRTAFVPLTADINESKKGGLG 381
Qy 410 MFLRAT-----EMNPSTINGCG--FPDVNRTEFKGLTTMTVTVLVLLQF 452
Db 382 VLRVSSRSRNCVEVERLLIQQMTTQTVALSGCKGFYFLRRLLFGMAGTIVTYEVLVLOF 439
```

```
RESULT 15
US-10-447-328-70
; Sequence 70, Application US/10447328
; Publication No. US20040003419A1
; GENERAL INFORMATION:
; APPLICANT: Carlson, John R.
; APPLICANT: Clyne, Peter J.
; APPLICANT: Marr, Coral G.
; APPLICANT: Yale University
; TITLE OR INVENTION: No. US20040003419A1el Taste Receptors in Drosophila
; FILE REFERENCE: 44574-5072
; CURRENT APPLICATION NUMBER: US/10/447, 328
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/09/593, 519
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138, 668
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 60/181, 704
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-447-328-70
```

```
Query Match 4.5%; Score 119; DB 15; Length 477;
Best Local Similarity 19.9%; Pred. No. 0.013;
Matches 101; Conservative 73; Mismatches 171; Indels 162; Gaps 22;
```

```
Qy 108 RIIGVLP-IVHGPRAKEFNMSASPIYSVVFVLLACYGVANNRIHYV-----RSL 160
Db 23 KIAGILPODLEKFSRNLKSRNGMIY-----MLSTLIILYVLYNLIYSFGEDRSL 76
Qy 161 SGPEEAVIAY--LFLVNLPIIMIPIILWYBARKIAKLFNDWDFEVLYYOI----- 210
Db 77 KA--SQSTLTFVIGLFTYIGLIMVY-----SDQLTALRNGRIGE--LYERIRLVDERL 127
Qy 211 --SGHSLEPLKARQKAVYAIYVLPILSVLVSVI--THVTMSD-----LNIQVVPYCI-- 258
Db 128 YKEGCVMDNSTTGRIRIMLTMTVIFELSLIVSTYKLVDSQMSLSLIVSAIPTFINT 187
Qy 259 LDNLTAMLGAMWFLICEAMSTIAHLAERFOKALKHIGPAAWADYRVLMRLSKLTRD 310
Db 188 LDKI-----WF-----AVSLYALKRFBAINATLEBIVDTHEKHKLMRLGNQEVPRP 234
Qy 311 -----LSKLTTRDTG-----NALC-----Y 324
Db 235 LDSQPPQYDSNLBYLYKELGMDIGSIGKSSVEBEKLNLCQVHDEICEIGKALNELWSY 294
Qy 325 TFFVMSLFLPITLITLISYGLMSQLSGFGIKDIGTLTALMNIGLLFYICDEAHYASVNV 384
```

```
Db 295 PILSIMAYGFLITTAQIYFL-----YCATQYQSIPSLFRSAKNPITVIYVL 340
Qy 385 RTNFQCKLMLVELNMNSDAQT-----EINMFLRATENPST--INC 424
Db 341 SYTSKCVVLIYLSMKTSQLSKRTGISLHKGVAADNLLYEIVNHLSLKLNHSVDPSA 400
Qy 425 CGPEPDVNRTEFKGLTTMTVTVLVLLQFOISIPDKDSGGANNITVVDVFMOSLDNDMS 484
Db 401 CGFETLMEETLYGVSGGITSYLLIILIQFNLAQAQAKAIGTFN-----SLNDTAG 450
Qy 485 LMGAST-----LSTTT---VGTTLPPPI 504
Db 451 LVGAATMDMNISSITLRDFTTTMTPAV 477
```

```
Search completed: May 23, 2005, 19:29:12
Job time : 144 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2005, 18:59:59 ; Search time 41 Seconds  
(without alignments)  
1201.536 Million cell updates/sec

Title: US-10-081-816-12

Perfect score: 2645  
Sequence: 1 MRPSGKVKVKGHGQNSGHS.....TTTVGTTLPPIPKLKGRKG 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

- 1: p1r1:\*
- 2: p1r2:\*
- 3: p1r3:\*
- 4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119.5	4.5	439	2 C22845	NADH2 dehydrogenas
2	118.5	4.5	789	2 T29461	hypothetical prote
3	112	4.2	443	2 F30010	NADH2 dehydrogenas
4	111.5	4.2	395	2 I64248	hypothetical prote
5	107.5	4.1	346	2 H82509	hypothetical prote
6	107.5	4.1	583	2 A11510	C-terminal domain
7	106.5	4.0	698	2 H98092	hypothetical prote
8	106.5	4.0	702	2 D95228	hypothetical prote
9	105.5	4.0	653	2 B71640	NADH2 dehydrogenas
10	105	4.0	453	2 H71732	Cytochrome P ubiqun
11	105	4.0	468	2 A75145	protein translocas
12	105	4.0	654	2 B64245	periplasmic phosph
13	105	4.0	766	2 B75059	probable transmemb
14	104.5	4.0	896	2 H86762	ABC transporter pe
15	104	3.9	369	1 G70007	conserved hypotet
16	103.5	3.9	317	2 T21047	NADH2 dehydrogenas
17	103.5	3.9	459	2 T17358	hypothetical prote
18	103.5	3.9	515	2 A02349	hypothetical prote
19	103.5	3.9	656	2 T31884	hypothetical prote
20	102.5	3.9	598	2 D71127	hypothetical prote
21	102.5	3.9	725	2 A90255	hypothetical prote
22	102	3.9	468	2 B71184	probable preprotei
23	102	3.9	925	2 A72096	ct34 hypothetical
24	102	3.9	925	2 B81573	conserved hypotet
25	102	3.9	925	2 B86527	CT34 hypothetical
26	101.5	3.8	1036	2 D70117	acetylflavine resis
27	101.5	3.8	437	2 S34959	NADH2 dehydrogenas
28	101.5	3.8	437	2 S73284	hypothetical prote
29	101.5	3.8	461	1 QXKL4M	NADH2 dehydrogenas

30	101.5	3.8	488	2 T20124	hypothetical prote
31	101.5	3.8	511	2 F84968	virulence factor m
32	101.5	3.8	891	2 A97718	hypothetical prote
33	101	3.8	459	2 A59154	NADH2 dehydrogenas
34	101	3.8	651	2 F90536	transport protein
35	100.5	3.8	467	2 F90546	probable terminal
36	100.5	3.8	580	2 A83874	Carbon starvation-
37	100.5	3.8	1483	2 S42839	thiazide-sensitive
38	100	3.8	379	2 PC4180	arsenical pump mem
39	100	3.8	429	1 C41903	NADH2 dehydrogenas
40	100	3.8	653	2 T14241	hypothetical prote
41	99.5	3.8	244	2 B64014	probable permease
42	99.5	3.8	301	2 G96944	glycerophosphoryl
43	99.5	3.8	583	2 AH1151	hypothetical prote
44	99	3.7	348	2 T21627	melanocortin recep
45	99	3.7	360	2 B46647	

ALIGNMENTS

RESULT 1

C22845 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Trypanosoma brucei mitochondrion  
C:Species: mitochondrion Trypanosoma brucei  
C:Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C:Accession: C22845  
R:Hemgens, L.A.M.; Brakenhoff, J.; De Vries, B.F.; Sloof, P.; Tromp, M.C.; Van Boom, J.  
Nucleic Acids Res. 12, 7327-7344, 1984  
A:Title: The sequence of the gene for cytochrome c oxidase subunit I, a frameshift contai  
ondrial maxi-circle DNA.  
A:Reference number: A93537; PMID:85037915; PMID:6093040  
A:Accession: C22845  
A:Molecule type: DNA  
A:Residues: 1-439 <HBN>  
A:Cross-references: UNIPROT.Q33575  
C:Genetics:  
A:Genetic code: SGC6  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match	Best local similarity	4.5%; Score 119.5; DB 2; Length 439;
Matches	71; Conservative	82; Mismatches 127; Indels 153; Gaps 17;
QY	105	WFLRIIGVLPVIRHGPAPAKFEMNSASPTYSVFFVLLACVGYVANNRHHVNSLSGPF 164
DB	51	WVFVFMGII-----MYLIFLSKKC---VSVMKVFYI----- 80
QY	165	EEAVIATVLF-VNLPIMIPILMYEARKLAKLFNDMD-----FEVLVYQISGSLP 216
DB	81	---VMIMYIYINVLIIIL-----DDFMCFMIAFSLFPJCLVSLF 120
QY	217	LKLRKAV-----YIAIVLPILSVLVTHVTMSD---LNIQVVPYICIDNLTAMLGAW 269
DB	121	FNFNRRPFAFYLLIRBSVSVCIICIVIHENININQAPIDVCYPSLSAIPW 180
QY	270	WFLICEAMSTHALLAERFQALRHIGPAAVADRYVWL-----RISKTR 316
DB	181	ILLF-----IMPAIKPIWPFH-----VMLDEMHEVNTENSVLASTVL 219
QY	317	DTGNALCTFPVMSL-----YLPFIITLSIYGLM-----SOLSEGGIKDIG 358
DB	220	KIGFGVYKFLFIAMNTISWFLGPTDSVIVLGLVFLAMSILFISDYKTIAMNSIHTG 279
QY	359	LTITALLNIGLLF-----YICDEAHYASV-----VRTNFOKKLMLVEMNNSDAQ 406
DB	280	IGLILMNDLILFLGILLICLHAILLSSFMFIYIGVYVNDVYGRIFLLISFGISIMS 339
QY	407	E-----INMFLRATENPSTINCGSPFDVNRILFKGLITMTVYLV 447
DB	340	SLFLCLFLFNIDPPFMLLFYVDIFLVGLISISFYIISFYIITLITLSI-----YIY 394

QY 448 VLLQFQISIPDTK 460  
 : : : : :  
 DB 395 MCLSFYSFWMDK 407

## RESULT 2

T29461

hypothetical protein T05B11.2 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29461

R:Geisel, C.; Bradshaw, H.

Submitted to the EMBL Data Library, March 1996

A:Description: The sequence of *C. elegans* cosmid T05B11.

A:Reference number: Z20621

A:Accession: T29461

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-789 &lt;GE1&gt;

A:Cross-references: UNIPROT:Q22211; EMBL:U53151; PIDN:AA37067.1; GSPDB:GN00023; CESP:T0

A:Experimental source: strain Bristol N2; clone T05B11

C:Genetics:

A:Gene: CESP:T05B11.2

A:Map position: 5

A:introns: 29/3; 123/2; 229/2; 344/1; 362/1; 418/3; 512/2; 618/2; 707/2

Query Match 4.5%; Score 118.5; DB 2; Length 789;

Best Local Similarity 20.8%; Pred. No. 0.14; Matches 100; Conservative 65; Mismatches 170; Indels 145; Gaps 20;

37 FLNAKPL-----NSANAQALYGVKRYSGLAERLDADYEAPLDRKSSDSTASNN 88

DB FNSNPDLVLSIILGFPLAIIPIFVLLIPCTLOEHL-----KSTNCLINN 277

QY 89 PEKK-----PSVF--YR-----NIDPIWFLRIGVLPVIRHGPAPAKFENASAFISYVF 138

DB 278 DEBECKTDPIVQYVYIMIDRLQKIEIRICIIASLPITONTFAACTFIIPNFGFLLIIRK 337

QY 139 FVLLACVGVVANNRRIHVSLSGPEBEAVIATLVNLIPIIPIWYEARKIATFN 198

DB 338 KYIL-----ANELDKNWKRNS-----LVFSEFDDKSELLG 370

QY 199 DWDDEFLVYIOISGHSPLKLRQAVYIAIVLPISVLSVITHVTMSDINOVVYCI 258

DB 371 PRRHF-----LKVSLDPSAKTKCCIFTCIV-----GVIIIVITPKMISVLMQAGAPFSF 422

QY 259 LNLITLMLG--AMWFLICE--AMSTIAHLAERFQALRHIGPAAVY-----DTRVLM 309

DB 423 SWGEASFEGYPAIFGFVCSLCLEFGWTKNGIASKFCRLVRVRLRQVAMPKLDNFRILHI 482

QY 310 RUSKLRDGNALCYTF-----VENSLLYFFIITLSY--GLMSQSLSEGGIGDITTA 363

DB 483 L-----ALCSIFPFAVMSWITITNFINGKIYYGIEONLSRIIFILVTNLYIW 532

QY 364 LWNIGLIFYICDEAHYASVNVRTNFOKLLMVELMNMNSDAQ----- 405

DB 533 ISTICAIIT-----FMSSALNREVYFNEBLQAKXKTKLKNIGVLEKED 578

QY 406 -----TEINMPLRATENPSTINCGFPDVNRTLFKGL-----TMTVYLVVL 449

DB 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

QY 579 FRONELIEMTLF-----ANGSLSLSGGFAPL--FLIMGLVNGIYLTNSFTTTPVLCIIL 631

A:Reference number: A22848; MUID:85079995; PMID:6096360

A:Accession: F30010

A:Molecule type: DNA

A:Residues: 1-443 &lt;DBL&gt;

A:Cross-references: GB:M10126

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC6

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C:Keyword: membrane-associated complex, mitochondrion; NAD; oxidative phosphorylation; c

Query Match 4.2%; Score 112; DB 2; Length 443;

Best Local Similarity 20.2%; Pred. No. 0.23; Matches 82; Conservative 65; Mismatches 131; Indels 128; Gaps 20;

97 YRNI--DPIN-WPLRIGVLPVIRHGPAPAKFENASAFISYVFPVLLACVGVYANNR 153

DB 43 YINIVYINLWFIYFMGLVFP-----LIFLSKLVSYSKYFIYL----- 85

QY 154 IHIVSLSGPFEAVIATLVNLIPIIPIWYEARKIATFNMDMD-----FEVL 206

DB 86 -----SYIFIFPDVTLIL-----IDDFMCFMILFESL 113

QY 207 YIQISGHSPLKLRQAVYIAIVLPISVLS-----VITHVTMSDINOVVYCI 257

DB 114 FEPICFVSLFENFNRRFPAIFLVFSSLSIMCIIICIIIFHNI--LNLQSPIDIC 171

QY 258 IIDNITLMLGAW-WPLICAMSTIAHLAERFQALRHIG-PAAMVADRYVLMRLSKLT 315

DB 172 IFDSL--YGLVYVWILFTMSI-----KPIPFHWMDELHAEVTELSILASV 222

QY 316 RDTGNALCYTFVMSL-----YLFPIITLSYGL-----MSQSLSEGGIGDI 357

DB 223 LKIGFGGLKFLFLSFGNSIMFLGIDSLVMLGLPLATLLPLSDYKIIATMSVHT 282

QY 358 GLTITLNM-----IGLIFYICDEAHYASVNVRTNFOKLLMVELMNMNSDAQTEINMFL 412

DB 283 GIGLILMINDIIFGLLIF--CNLSHITS-----SAFMEMMGVMDNYGVRIFLML 333

QY 413 RATEMNPSTINCGFPEDVN--RTLFGKLLTWTYLVVLLQFOIS 456

DB 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

QY 334 IS-----PFGISIMSSLGIFLFINIDPFML--FYIDI 366

[illegible]

Db 297 LVVIMMN-----LARFLGLAGLEKDTLVLFGLTTLTVLVVKGLERTPLXKHL 344

RESULT 6

A11510

C:Terminal domain glycerophosphoryl diester phosphodiesterase homolog lin0625 [imported]

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C:Accession: A11510

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Eutican, K.D.; Fahl, H. D.; Jones, L.M.; Karet, U.

Science 294, 849-852, 2001

A:Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Matcounam, A.; Matok, C.; Schuierer, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

Article: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: A11510

A:Status: preliminary

A:Molecule type: DNA

A:Releases: 1-583 <GLA>

A:Cross-references: UNIPROT:Q92E36; GB:AL592022; PIDN:CA695857.1; PID:916413065; GSPDB:GN

C:Genetics:

A:Experimental source: strain Clijp11262

A:Gene: lin0625

Query Match 4.1%; Score 107.5; DB 2; Length 583;

Best Local Similarity 21.4%; Pred. No. 0.76;

Matches 84; Conservative 66; Mismatches 104; Indels 139; Gaps 24;

QY 106 FLRTIGVLPYVH---GPRAKFEMNSAFIYSVVFVLLACVGVYANNRHHVRLSG 162

DB 18 YLKVTLTLTIOAFYIGP-----FYIFFFPLIKVIGVPGITDNL-----G 59

QY 163 PEEBVIATLFLVNIPLPIWPIPLWYKARKAKLFNDWDFEVL--YYQISGSLPL--- 217

DB 60 EYFSSPVAVIML-ILALILFLVYEL-----GFFIMAIYQLRGSEYTVFKI 107

QY 218 --KLRQKA-----VYIAVLPLISLVSVITHVMSDLNINOVVPCILDNL- 262

DB 108 IORLVKAKYFLSYQAIYFLVFLYFLLPLIAG-LSLPIIT-----ITENLYLPHRTIDELM 160

QY 263 TAMLGAWFLIEAMSTHMLAERQKALKHIGPAAVADYVLMRLSKLRDTGNAL 322

DB 161 KTTGTWLVVI--ATAIIFYISA-RLVFALPY----FLRDKS--LKIS-----GAI 202

QY 323 CTFEPMSLYLF--TITLSIYG-----LMSQSEGGIKDGLTIT 362

DB 203 RKSMPFSQHLPEFLMKLVIIIVIGFLVSIITITIMPLLVBKIRPGIAVIAGITLT 262

QY 363 ALMNIGLL-----FYICDEAHYASVNVATNF--QKLLMVELIMNNSD 403

DB 263 ILQVIGFPAAGIFQGIILAQLVGNAPAI--EGGHAPV-ANQPMPKKRRFIIV----- 311

QY 404 AQTETNMFLKATEMNPSTINCQGFEDVARTLEK 436

DB 312 --AIIVFIIFSGFN-----IFTVNATLYE 333

RESULT 7

D95228

hypothetical protein SPI953 [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: *Streptococcus pneumoniae*

C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C:Accession: D95228

R:Retselein, H.; Neilson, K.E.; Paulsen, I.T.; Bisen, J.A.; Reed, T.D.; Peterson, S.; Heidon, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple, neon, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.

Article: Complete Genome Sequence of a Virulent isolate of *Streptococcus pneumoniae*.

A:Reference number: A95000; MUID:21537209; PMID:11463916



A/Accession: D95228  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-698 <KUR>  
A/Cross-references: UNIPROT:Q97NR4; GB:AE005672; PIDN:NAK76021.1; PID:G14973459; GSPDB:G.C  
A/Experimental source: strain TIGR4  
C/Genetics:  
A/Gene: SPI953

```

Query March 4.0*: Score 106.5; DB 2; Length 699;
Best Local Similarity 19.1%; Pred. No. 1.1;
Matches 128; Conservative 84; Mismatches 175; Indels 283; Gaps 30;

QY 6 EKVVKGH-----GQSGSHSLSGANVYRRKKGDVFLNAKPLPNSANAQAAYLVGVRYS 59
DB 88 EKISKNNFYILDPAKGQGRMSISEFERHYSN-----ILLTFKKLDS-----FMSK-- 133
QY 60 IGLAERLDADYEAPPLDRKSSDSTASNPEFKSVFRNIDPINMFLR-----IIGVLP 114
DB 134 -----DNKKS-----PYLAKFFPKYRNKGLIFVLTALLVYIGSLVP 169
QY 115 IVRHGPARAKEMNSAS-----FIYSVVF-----FYLACVYG----- 147
DB 170 IANRYIIDYFNKODSYSSRMLFTLLFPTVSEFSMLYLRQIYASLKYIMDKESISYDFMK 229
QY 148 -----YVANNRIHVRSLSGPFEAAVAYLP----- 172
DB 230 HLILPYSFYEKRTGLDILFRANSIVYIREILSNFNFAIILDLMIYVAVVLEPSKFM 289
QY 173 --FLVNI--LPIMIPILMYEAARKIAK-----LF 197
DB 290 VIFISLISLALSIWYPIIKISKMLIDKNIEKYNVQNTSEVISKNSDIKLTGEEEFWI 349
QY 198 NDMDPEFVLYYQISGSHLPLKROKAYIAVLPILSVLSVYITHNM-----SDL 248
DB 350 NKMNENFTKOLIIG-----RKLDIHLSTSVSSITVNLQILPLVLTLIVGVNIKTEPQL 401
QY 249 NINQV-----PYCI-----LDN--LTAMLGAMFPL--ICEAMSITLHLAERFOKALKH 294
DB 402 TLGGIVASIVSTWSPFISPIIISLSDNYIQMLLKGYFLARIEVFNPKSEILPERISOIKF 461
QY 295 IGPAAWADYRV---LWLRLSKLTRDTGNALCYTFVFMSLYLPFIITLSIYGLMSQLSE 350
DB 462 -----DKKIELKDIWYK-----YGLF---D 478
QY 351 GFGIKDIGLT-----TALNMIGLIFYICDEAHYASVNV----- 384
DB 479 DYVLKGINVITIKKGETVAIVGESGSGSKYLAIKILLGLL-----EPNIGSIEVDGVEKEEI 533
QY 385 -RTNFOKKLMLVELLNMMNSDAQTEINM---FLRATEMNPSTINGCGPFVDVNRILPFKGL 439
DB 534 GQTIYRKIFQAVLONSTLSTYGTALRENLTFGHFVSDDEL--WNNLNSIGLSNVVSLPLGLE 592
QY 440 TTMV-----TYLVLLIQFOISLPTDKGDSGANNTIVDFVWDSLDNDNSLNGAS 489
DB 593 TYIAEGNNSFGQGOQOIMLLARCLLSKP-----SVVLDEATSSIDLNLISQQLTTS 642
QY 490 TLTSTTVGTT 499
DB 643 YLS--EIGTF 650

```

## RESULT 8

hypothetical protein clyH [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: H98092  
R:Hoeklin, U.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; Ely, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.  
Y: Bacteriol. 183, 5709-5711, 2001  
Y:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balz, R.H.; Jaskunas, S.R.

A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6  
A;Reference number: A97872; MUID:21423245; PMID:11544234

A:Accession: H98092  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-702 <KUR>  
A:Cross-References: UNIPROT:Q8DNF4, GB:AE007317, PDB:5I5459452, GSPDB:GN  
C:Genetics:  
A:Gene: clyB

Query Match	Similarity	4.0%	Score 106.5	DB 2	Length 702
Beet	Local	19.1%	Pred. No. 1.1	175	Indels 283
Matches	128	Conservative	84	Mismatches	30
QY	6	EKVNVGH-----	GCGNSGSHLSGMANYRRKKGDAVFLNAKPLNSAQAAYLVGYAKYS	59	
DB	92	EKISNKKFYIIDPARGROMSISEPERHYSN-----	IIILFKLDS-----	FMSRK--	137
QY	60	IGLAERLDADYEAPPLDRKSSDSTASNNPEKPSVYRNIDINMPLR-----	IIGVLP	114	
DB	138	-----	DNKKS-----	PVLKPYFKRYENKJGILFEFVALLVYIQSLVP	173
QY	115	IYRHGPAPAKPEMNSAS-----	FYSVVF-----	FVLACYG	147
DB	174	IANYIIDTNEFKDYSRMLFTLIFTVSSLMALIKQIYASLKYIMDEISYDFMK	233		
QY	148	-----	YVANNRHHVRSLSGPEEAVIAYL-----	172	
DB	234	HLIYLPYSFEKRTGDLIFRANSIYVIREILSNFIALLIDLLMIVVAVLVESFKM	293		
QY	173	--FLVNI--	LPIMTIPILMYEARKIAK--	197	
DB	294	VIFLISLALSIWVYPIIKISKNLIDKNIKEKVNYQITSEVISKNSDIKLGEBEFWI	353		
QY	198	NDMDPEVLVYQISGHSPLKLRQKAVYTAIYLPILSYLSVVTHTM-----	SDL	248	
DB	354	NKMDENFTKQILIG-----	RKLDHLSIVSSITNVYQIILPVLTILVGNIKTFEQL	405	
QY	249	NINQNV-----	PYCI-----	LDN-LTAMLGAMWPL-ICEAMSIYAHLLAERFOKALKH	294
DB	406	TLGQVVALSTVSPYFISPLISLDNYIQMLKGYTLREDVNTSSELPERVSDIKP	465		
QY	295	IGPAAVADYRV-----	LWLRLSKLTRDTGNALCYTFVMSLYLFTIITSIYGLMSQLSE	350	
DB	466	-----	DKRIELDKIWIYK-----	YGLF-----	482
QY	351	GGGIDIGLTI-----	TALMNLGLFPCDBAHYASVNV-----	384	
DB	483	DYVNLGINVITLKKGTVAIVGSSGSKSTLAKILIGLL-----	EPNIGSIEVGVEXKEI	537	
QY	385	-RTNFQKLLMWELWMNSDAQTEINM-----	FLATENNPSTINCSPFDVNRITLFFKGL	439	
DB	538	GQTLYRKIFGAVLQNSTLSYGTLRKNTLGHFVSDSEL-MTNLNSIGLSNVVKSPLPGLE	596		
QY	440	TTMV-----	TYLVLLQFOISIPFDKDSGEGANNITVVDYVWDSIDNMSLNGAS	489	
DB	597	TIABEGNNSGGQOQMILLARCLSKP-----	SVVVLDEATSSLDNLSQQLTTS	646	
QY	490	TLSTTVGTT	499		
DB	647	YLS--EIGTT	654		

## RESULT 9

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain L1 RP792 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 09-Jul-2004  
C:Accession: B71640  
R:Anderson, S. G. E.; Zomorodipour, A.; Andersson, J. O.; Sichelitz-Ponten, T.; Almarik, U.  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
#:Reference number: AV1630; MUID:J9039499; PMID:9823893

QY 96 FYRNIDINWFLRIGVLPVRHG--PARAKFENNSAFIYSVFEVLLACVYGIVANNR 153

periplasmic phosphatase permease homolog - *Mycoplasma genitalium*  
C;Species: *Mycoplasma genitalium*

C>Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: E64245  
 A:Cross-references: UniProt:P47651; GB:U39726; GB:L43967; NID:g1046113; PID:g1046125; TI  
 A:Experimental source: strain G-37  
 C/Genetics:  
 A:Genetic code: GCG  
 A:Start codon: GTC  
 C:Superfamily: periplasmic phosphate permease Ag88

Query Match 4.0%; Score 105; DB 2; Length 654;  
 Best Local Similarity 21.3%; Pred. No. 1.4;  
 Matches 104; Conservative 73; Mismatches 153; Indels 158; Gaps 27;

QY 89 PERKSVFPRNIDPIN-----WFLRIIGVLPVHGPAPAKPEMNSASPIYSVFVLLA 143  
 DB 46 PDFAKSLFNLFFNLGNKQAGIWP-----PLT-----VSPTVS-IGALITA 84  
 QY 144 CYVG-----YVANNRIH-----IVRSLSGPFEEAVIAYLFLVNIPLMIPIIPLYEA 190  
 DB 85 SYIGVARTSPLYRCKPKRKLSLIIIDLSG--IFSALFGLPASQILSIF----- 134  
 QY 191 RKIAKLFNDWDPEVLVYQISGHSPLKLRKAVYA-IVLPTL-----SVL 236  
 DB 135 RLILKL-----PPLSLNLVIALMSFMIIPIVISLTNTLTLYVNDI 176  
 QY 237 SVVITHTVMSDANINOVPCILNITAMLG-AWMLICGMSITAHLAERQKALKKI 295  
 DB 177 SVVSLGKNTSAIYKIKKEIKPQLTVLITLAFARAISETMAVNVLOSVAHQEVINN- 235  
 QY 296 GPAAMADVRLMLRSLK-LTRDTG-----NALCYTF-----VEMSLYLFPII-----TL 339  
 DB 236 -NRFLSDAKTIGSVSTIFSENGDQNGVYITGIIILIVSLNFAIWSANPKTL 294  
 QY 340 SYIGLMSQISEGFGIKDIGITITAM-----NIGLAFYICDEAHYA-----SVNRTNF 388  
 DB 295 ERYPLFKKIS-----NFIYQVWFIPNNISALFVLDLSTFQSVKIKVNNINERSLP 346  
 QY 389 OKKLLM-----VELNMNSDAQTEINMFLR-----ATEMNPSTINGCGF 427  
 DB 347 FFERLQSVVWIKLNYFLKIFQELICTFLAFGFVALIILFVFIINGSVAINNNGSTVPS-- 403  
 QY 428 FVNRTLFKGLTTLTVTVLVLLQFQISIP-----TDKQDSEGANNTVVDVPMDS 478  
 DB 404 FEARST-GALVNTLV---IILITITITFPALLIIMALEYNNKSVKVN--VFENVIDS 457  
 QY 479 LNDMSLM 486  
 DB 458 LSSMPSII 465

## RESULT 13

B75059  
 probable transmembrane oligosaccharyl transferase PAB0974 - *Pyrococcus abyssi* (strain OX  
 C/Species: *Pyrococcus abyssi*  
 C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C/Accession: B75059  
 R:Anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru  
 A:Reference number: A75001  
 A:Accession: B75059  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-766 <RAW>  
 A:Cross-references: UniProt:Q9UYP5; GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB5036  
 A:Experimental source: strain Oresay  
 C/Genetics:  
 A:Gene: PAB0974

Query Match 4.0%; Score 105; DB 2; Length 766;  
 Best Local Similarity 20.1%; Pred. No. 1.7;  
 Matches 111; Conservative 70; Mismatches 188; Indels 182; Gaps 24;

QY 18 GHSLSGMANYRRKKGDVAFLNAKPLNSNAQAQILYGVKYSIGLAERLD-----ADYEA 72  
 DB 139 GHIFRSMANYR--GD-----NYMLFWYSIALAGIAYALRTKGYRRLLVFYAV 183  
 QY 73 PLIDRKSSDSTASNNPEFKSVFERNIDPINMFLRIIGV-LPIVHGPAPAKPEMNSAS 131  
 DB 184 PTLASGIS-----SYFMQATYPLVFLSLNGVFLAISGLADKRNFLDSFT 230  
 QY 132 FLYSVFVFLACYV-----GYVANNRIHVSLSGPFEEAVIAYLF--LVNIPLIM-- 181  
 DB 231 IILSTAFGAILIANYLGEKFGYGMIGYNRQHIYSKIGIKLRDAYLFIHLVLPISLG 290  
 QY 182 IIPILMYEAR--KIAK----- 196  
 DB 291 LIIILFLSRFPYKSKAKAGIIVIGLTVSILILKLPALRGLGIFDMFKSTPIMETRP 350  
 QY 197 --FND-WDDFEV-----LVY----- 208  
 DB 351 TTFHDMKAFSISIFLPLFLRHFPEKKTEDFLGLGIYVSLYMLAMAFVFGSLA 410  
 QY 209 -----QISGHSPLKLRQ--KAVYIAIVLPISLVSIVTHTVMSDL-----NINOV 253  
 DB 411 VATMAGIGLVEGSLVYIQRRKSGKASRVALLIIL--LIIVNGAFPLKQCLSMRPLINKE 468  
 QY 254 VEPYCL-----DNLTFMLGAWFPL-----ICEMSTAHLAERQKALK 293  
 DB 469 WENALTLWLNNSNENDVILLAMWDGAMITTYSRAPVAEIAENPDVALYLLGARNDWM 528  
 QY 294 HIGPAMVADVRLMLRSLKLTREDGNALCYTFEVMSLYLFPIITLSYIGLMSQISEGFG 353  
 DB 529 SLGVDVIVSY-YDPLAKFSIVRTASQSGVNLKRYIAVLPLSSYGGIL--IFEGGE 585  
 QY 354 IDDIGLTTALNNIGLFTICDEAHYAVNRTNFQKLLMYELNMNSDAQTEINM-FL 412  
 DB 586 YKIIAKP-GDIWDVRII--IGDHVYVPR-GLVVEYKGVTESKLTAKSYSDAYLYINLAK 641  
 QY 413 RATEMNPSTIN 423  
 DB 642 YAILMNSSTFN 652

## RESULT 14

H86762  
 ABC transporter permease protein ylbB (imported) - *Lactococcus lactis* subsp. *lactis* (str  
 C/Species: *Lactococcus lactis* subsp. *lactis*  
 C/Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
 C/Accession: H86762  
 R:Boletcin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlic  
 Genome Ref. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssi  
 A:Reference number: AB6625; MUID:21235186; PMID:11337471  
 A:Accession: H86762  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-896 <STO>  
 A:Cross-references: UniProt:Q9CGJ2; GB:AE005176; PID:g12724061; PIDN:AAK05202.1; GSPDB:GT  
 A:Experimental source: strain ILL403  
 C/Genetics:  
 A:Gene: ylbB

Query Match 4.0%; Score 104.5; DB 2; Length 896;  
 Best Local Similarity 22.3%; Pred. No. 2.3; Indels 67; Gaps 13;  
 Matches 63; Conservative 42; Mismatches 111;

QY 157 VRLSGPFEBAVAYLFLVNLIPM-----IIPILMYARKIADKLFNDM-----D 201  
DB 361 IDLSISNIFPVLLFPAIALVSLTMTFRPVEBERGNGLKALGYSNRDIRKKFMYGLVSS 420  
QY 202 DPEFLVYQISGHSPLKLRQKAVYIAVLPILSVSVVITHVMSDININ-----QVVPY 256  
DB 421 GLGALVTTIIGHTF-----LPI-AVFNAVITASSTSNRLTSLPMTIVAF 465  
QY 257 CIIDNLTAMLGAMWFLICEAMSTIAHLAERFQALKHIGPAMVADYRLMTLSKLTR 316  
DB 466 AIA-IACSLLPAYVAVVMBELKEVPASLFLAKVPFA-----GSRILBKINFIWKMSFTYK 520  
QY 317 DTGNALCTYVFNLSLYFFIITLSIYGLMSQLSGFGIKDIGLITTLAMN--IGLLYYIC 374  
DB 521 VTARNL---FRYKKRMIMTI--FGVAGCTALLVVGFGRD---SIGLSNKGQOIL--- 569  
QY 375 DEAHYASVNRVTNFKKLLMVELNMNSDAQTEINMFLRATEN 417  
DB 570 ---HY-----DMITTEKKKVDKKEKIDKLASSEI 598

## RESULT 15

G700007

conserved hypothetical protein yuef - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C/Accession: G700007

R.;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez  
C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, A.; Lapidus, A.; Lardinois,  
Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Seanton,  
A/Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serot  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A/Authors: Yoshikawa, H.P.; Zumschein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A/Reference number: A69580; MUID:96044033; PMID:9384377

A/Accession: G700007

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Cross-references: UNIPROT:O32095; GB:Z99120; GB:AL009126; NID:92635613; PIDN:CAB15168.

A/Experimental source: strain 168

C/Genetics:

A/Gene: yuef

C/Superfamily: Bacillus subtilis conserved hypothetical protein yuef

Query Match 3.9%; Score 104; DB 1; Length 369;

Best Local Similarity 18.7%; Pred. No. 0.83;

Matches 83; Conservative 65; Mismatches 139; Indels 156; Gaps 19;

QY 105 WFLRIIGLTVIRGPARAKPEMNSASFIY-----SVPEFVLACVGVANNRI--- 154  
DB 9 WTLQILFVLLI-----FVATKVSFVQPFIVFISTLFFPMILAGILYFIRPVRL 60  
QY 155 ---HIVSLSGPFEBAVAYLFLVNLIPM---IIPILMYARKIADKLFNDMDF---E 204  
DB 61 LEKIPRTLS-----ILIIYLLFGLAFISAVGPIL---TAQVTGLENMLPDYIKQIQ 112  
QY 205 VLYVOISGHSPLKLRQKAVYIAVLPILSVSVVITHVMSD 247  
DB 113 ALTYKOLS-HSQWFTWMNQDYSISKIQLSTPLQNLPTNTSLSAVFGVNTITLVI 171  
QY 248 LINVGVVPCIDNLTAMLGAMWFLICEAMSTIAHLAERFQALKHIGPAMVADYRL 307  
DB 172 ITVPILFYMVKD-----GHRF-----PHLAVKILPASRYTE 203

QY 308 WRLSKLTRPT-----GNALCTYVFNLSLYFFIITLSIYGLMSQLSGFGIKDIGLTI 361  
DB 204 GLKIFKQLSDTLAAYRQGLLILCFVGTACFISGLNGLPYALI----- 247  
QY 362 TALNIGLIFYIDEAHYASVNRVTNFKKLLMVELNMNSDAQTEINMFLRATENPST 421  
DB 248 ---LGIWALINIIY-----VGPIGAA---PAV 271  
QY 422 INCGFPDVNRITLFGKLLTMTVTVLVLOFQISIPDKDSEGANNTVVDPYMDSLND 481  
DB 272 I--VGFMV---SPAKLFAIIVVIVQOLDGNLSPVIGKRLNTHPLTII----- 317  
QY 482 DMSIGASTLSTTTCVGTLPPEI 504  
DB 318 -LILIGAGSFG-GILGMITLAVPV 338

Search completed: May 23, 2005, 19:25:56

Job time : 44 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2005, 18:58:23 ; Search time 171 Seconds  
(Without alignments)  
1533.243 Million cell updates/sec

Title: US-10-081-816-12

Perfect score: 2645  
Sequence: 1 MRPSGKRVKKGQNSGHS.....TTTVGTTLPPIIMKLKRRKG 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2645	100.0	512	1 G63A.DROME	O9VZL7 drosophila
2	1276	48.2	379	2 Q7PYF4	O7PYF4 anopheles g
3	480	18.1	485	2 Q7PQT4	O7PQT4 anopheles g
4	458	17.3	467	2 Q7PMG3	O7PMG3 anopheles g
5	405	15.3	447	1 G21A.DROME	O9VPT1 drosophila
6	167.5	6.3	412	1 G98D.DROME	O83M5 drosophila
7	149.5	5.7	383	2 Q7PVK1	O7PVK1 anopheles g
8	148.5	5.6	472	2 Q7PEV2	O7PEV2 anopheles g
9	145.5	5.5	404	2 Q7PVJ9	O7PVJ9 anopheles g
10	143	5.4	444	2 Q7P85	O7P85 anopheles g
11	142	5.4	416	2 Q7PK14	O7PK14 anopheles g
12	141.5	5.3	411	2 Q7PVK0	O7PVK0 anopheles g
13	140	5.3	382	2 Q7PVU8	O7PVU8 anopheles g
14	139.5	5.3	451	2 Q7PKC5	O7PKC5 anopheles g
15	139.5	5.3	1210	2 Q7OKL0	O7OKL0 anopheles g
16	135.5	5.1	369	1 G39B.DROME	P58960 drosophila
17	132.5	5.0	406	1 G64B.DROME	P58956 drosophila
18	130.5	4.9	372	1 G391.DROME	P58956 drosophila
19	130.5	4.9	393	2 Q7PKA1	O7PKA1 anopheles g
20	130	4.9	403	2 Q7PK35	O7PK35 anopheles g
21	129	4.9	436	1 G61A.DROME	O9W0M2 drosophila
22	127	4.8	388	2 Q7PK39	O7PK39 anopheles g
23	126	4.8	390	2 O8REF8	O8REF8 fusobacteri
24	126	4.8	396	2 Q7PK37	O7PK37 anopheles g
25	125.5	4.7	386	1 G22B.DROME	P64180 drosophila
26	125	4.7	399	2 Q7PK31	O7PK31 anopheles g
27	125	4.7	403	1 G98B.DROME	O9V826 drosophila
28	125	4.7	451	1 G64E.DROME	P83296 drosophila
29	125	4.7	460	2 Q7KVS3	O7KVS3 drosophila
30	124	4.7	401	2 O8WMT7	O8WMT7 heliothis v
31	124	4.7	419	2 Q9WDS2	Q9WDS2 phytonomae

32	123.5	4.7	400	2 Q7PK34	O7PK34 anopheles g
33	123	4.7	405	2 Q7PK33	O7PK33 anopheles g
34	122.5	4.6	404	2 Q7PKT5	O7PKT5 anopheles g
35	122	4.6	414	1 G02A.DROME	Q9W54 drosophila
36	121.5	4.6	456	1 G64A.DROME	P83293 drosophila
37	121	4.6	397	2 Q7PFJ8	O7PFJ8 anopheles g
38	120.5	4.6	399	2 Q7PK38	O7PK38 anopheles g
39	120	4.5	384	2 O83X6	O83X6 enterococcu
40	118.5	4.5	4099	2 Q7RPL9	O7RPL9 plasmodium
41	118	4.5	396	2 Q7PK36	O7PK36 anopheles g
42	117.5	4.4	445	2 Q7PE20	O7PE20 anopheles g
43	117	4.4	429	2 Q7P194	O7P194 anopheles g
44	117	4.4	572	2 O6JV99	O6JV99 bombyx mand
45	116	4.4	397	2 Q7PVJ7	O7PVJ7 anopheles g

## ALIGNMENTS

RESULT 1  
ID G63A.DROME STANDARD; PRT; 512 AA.  
AC O9VZL7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DB Putative gustatory receptor 63a.  
GN Name=G63a; ORFNames=CG14979;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=2019606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Chertey J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunn P.,  
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,  
RA Glisler A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibbegan C.,  
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasero P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M., A.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,  
RA Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler B., Spreading A.C., Stapleton M., Strong R., Sun B.,  
RA Svrtkals R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";   
 RL Science 287:2185-2195(2000).   
 RN [2]   
 RP GENOME REANNOTATION.   
 RX MEDLINE=22426069; PubMed=12537572;   
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,   
 Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,   
 Smith C.D., Tang J.L., Whitefield E.J., Bayraktaroglu L., Berman B.P.,   
 Beitzencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,   
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,   
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,   
 Lewis S.E.;   
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a   
 RT systematic review";   
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).   
 RN [3]   
 RP IDENTIFICATION.   
 RX MEDLINE=21407712; PubMed=11516643; DOI=10.1016/S0960-9822(01)00258-5;   
 RA Dunipace L., Meister S., McNealy C., Amrein H.;   
 RT "Spatially restricted expression of candidate taste receptors in the   
 RT *Drosophila* gustatory system.";   
 RL Curr. Biol. 11:822-835(2001).   
 CC -1- FUNCTION: Probable role in the gustatory response.   
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).   
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor Dr-tr   
 CC family.   
 -----   
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration   
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-   
 CC the European Bioinformatics Institute. There are no restrictions on its   
 CC use by non-profit institutions as long as its content is in no way   
 CC modified and this statement is not removed. Usage by and for commercial   
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>   
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).   
 -----   
 CC EMBL: AEO03479; AAF47803.1; -   
 DR FlyBase; Fggn0035468; G63a.   
 DR GO; GO:0016021; C:integral to membrane; NAS.   
 DR GO; GO:0006527; C:platelet receptor activity; NAS.   
 DR GO; GO:0050912; P:perception of taste, sensory transduction o. . .; NAS.   
 KW G-protein coupled receptor; Glycoprotein; Hypothetical protein;   
 KW Multigene family; Receptor; Transmembrane.   
 FT DOMAIN 1 129   
 FT TRANSSEM 130 150   
 FT DOMAIN 151 166   
 FT TRANSSEM 167 187   
 FT DOMAIN 188 222   
 FT TRANSSEM 223 243   
 FT TRANSSEM 244 252   
 FT TRANSSEM 253 273   
 FT TRANSSEM 274 324   
 FT TRANSSEM 325 345   
 FT DOMAIN 346 350   
 FT TRANSSEM 351 371   
 FT DOMAIN 372 436   
 FT TRANSSEM 437 457   
 FT TRANSSEM 458 512   
 FT CARBOHYD 431 431   
 FT CARBOHYD 431 431   
 SQ SEQUENCE 512 AA; 57465 MW; B98BD94D6DABFF66 CRC64;   
 Query Match 100.0%; Score 2645; DB 1; Length 512;   
 Best Local Similarity 100.0%; Pred. No. 5e-189;   
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;   
 QY 1 MPSPGKVVYKGGQSGHSLGSMANYTRRKGDVAFVLAANKPLNSANAQALYGVKRYKSI 60   
 DB 1 MPSPGKVVYKGGQSGHSLGSMANYTRRKGDVAFVLAANKPLNSANAQALYGVKRYKSI 60   
 QY 61 GLAERLDADYEAAPLRLKSSSDTASNNPEFKSPVYRNIDPINFELRIIGVPIYRHGG 120   
 DB 61 GLAERLDADYEAAPLRLKSSSDTASNNPEFKSPVYRNIDPINFELRIIGVPIYRHGG 120

QY 121 ARAKPEMNSAFIYSVVFVLLACVGYVANNRIHVRSLSGPFEEAVIAYFLVNLPI 180   
 DB 121 ARAKPEMNSAFIYSVVFVLLACVGYVANNRIHVRSLSGPFEEAVIAYFLVNLPI 180   
 QY 181 MIPIIWEARKIAKLFNDWDFEVLVYQISGHSPLKQKAVYAIYVLPISVLSVVI 240   
 DB 181 MIPIIWEARKIAKLFNDWDFEVLVYQISGHSPLKQKAVYAIYVLPISVLSVVI 240   
 QY 241 THTVMSDLINOVVPCILNDITAMGAMPFICEMSTIHTHLARPOKALKHIGPAM 300   
 DB 241 THTVMSDLINOVVPCILNDITAMGAMPFICEMSTIHTHLARPOKALKHIGPAM 300   
 QY 301 VADRYLMLRLSKLTREDTGNALCYTFEWSLYLFFITTSIYGLMSQSEGFQIXDGLT 360   
 DB 301 VADRYLMLRLSKLTREDTGNALCYTFEWSLYLFFITTSIYGLMSQSEGFQIXDGLT 360   
 QY 361 ITPLMNIIGLLFYICDEAHYASVVRNTPQKILMVELMNMNSDAQTEIMFPLATEMNS 420   
 DB 361 ITPLMNIIGLLFYICDEAHYASVVRNTPQKILMVELMNMNSDAQTEIMFPLATEMNS 420   
 QY 421 TINCGEFDPVNRPLPKGLITWTYLVLLQFOISIPDKGSEGANNTIVDFVMSLD 480   
 DB 421 TINCGEFDPVNRPLPKGLITWTYLVLLQFOISIPDKGSEGANNTIVDFVMSLD 480   
 QY 481 NDMSLMGASTLSTTVGTTLPPIMLKGKRG 512   
 DB 481 NDMSLMGASTLSTTVGTTLPPIMLKGKRG 512

## RESULT 2

## ID Q7PYF4

PRELIMINARY; PRT; 379 AA.

## AC Q7PYF4;

DT 01-MAR-2004 (T-EMBLrel. 26, Created)

## DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)

DE 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

## DE AGCP12374 (Fragment)

GN Name=agCG53552; ORFNames=ENSGAG0000017642;

## OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

## OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

OX NCBI\_TaxId=180454;

## RN [1]

RP SEQUENCE FROM N.A.

## RC STRAIN=PEST;

RA Anopheles Genome Sequencing Consortium;

## RU Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an   
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is   
 CC preliminary data.   
 DR EMBL; AAAB01008987; EAA01019.1; -   
 DR NON TER 1 1   
 FT SEQUENCE 379 AA; 42958 MW; 135438038754E0F2 CRC64;

## Query Match

Best Local Similarity 48.2%; Score 1276; DB 2; Length 379;   
 Matches 241; Conservative 50; Mismatches 72; Indels 0; Gaps 0;

## QY 97 YNIDPIMFLIIGLPIYVRGPARAKFEMNSASFIYSVVFVLLACVGYVANNRIHI 156

DB 97 YNIDPIMFLIIGLPIYVRGPARAKFEMNSASFIYSVVFVLLACVGYVANNRIHI 156

## QY 157 VSLSGPFEEAVIAYFLVNLPIPIIMTILMWEARKIAKLFNDWDFEVLVYQISGHSPL 216

DB 157 VSLSGPFEEAVIAYFLVNLPIPIIMTILMWEARKIAKLFNDWDFEVLVYQISGHSPL 216

## QY 217 LKLRQAVYAIYVLPISVLSVVI 276

DB 217 LKLRQAVYAIYVLPISVLSVVI 276

## QY 277 MSITLHLLAERFOKALKHIGPAMVADRYVLMRLSKLTREDTGNALCYTFEWSLYLFFI 336

DB 277 MSITLHLLAERFOKALKHIGPAMVADRYVLMRLSKLTREDTGNALCYTFEWSLYLFFI 336

## QY 336 184 LSTAKTILAEDEFORALRHVGPAAKVEYKSLMLRLSKLARDGFTCYTFETICLYLFEI 243

DB 336 184 LSTAKTILAEDEFORALRHVGPAAKVEYKSLMLRLSKLARDGFTCYTFETICLYLFEI 243



QY 337 ITLSIYGLMSQSGFGIKDITLITLALMNIGLFIYCDEAHYASVNRINPCKLLMVE 396  
 DB 244 ITLSIYGLMSQISGDFGVKDIGLAVTAFCSVGLFIYICDEAHYASVNRINPCKLLMVE 303  
 QY 397 LMMNSDAQREINNFLEAETENPSTINGCFEDVNRITLFGKLLTMTTYLVLLQFOISI 456  
 DB 304 LMMNSDAQREINNFLEAETENPSTINGCFEDVNRITLFGKLLTMTTYLVLLQFOISI 363  
 QY 457 PTD 459  
 DB 364 PDE 366

## RESULT 3

Q7PQT4 PRELIMINARY; PRT; 485 AA.  
 AC Q7PQT4; 01-MAR-2004 (Tremblrel. 26, Created)  
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)  
 DE ENSANGP0000003178.  
 GN Name=ENSANGG00000002573;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 RN NCBI\_TaxID=180454;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -! CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAAB01008879; EMBL08342.2; -  
 SQ SEQUENCE 485 AA; 55676 MW; 8F30114775BFF5BF CRC64;

Query Match 18.1%; Score 480; DB 2; Length 485;  
 Best Local Similarity 29.6%; Pred. No. 1.8e-27;  
 Matches 113; Conservative 93; Mismatches 170; Indels 6; Gaps 4;

QY 96 FYRNIDPIMFLRIIGVLPYRHGPAKAFEMNSAFIYSVFPVLLACVGYANNRH 155  
 DB 102 FYRDKKLLVLFRLGAVMPITRSVGRITFSMRSAASYACFYLVSTVILVVGEXIK 161  
 QY 156 IVRSLSGFEERAVLAVFLVNLPIIMIPIL-WYEARIAKLFDMDFEVLXYOISGH 214  
 DB 162 VFQTTT-KEDYIYIGILVFVLPVHPFWLPVGMGVAKQVALYKTMGAFQRYIRVGT 220  
 QY 215 LPL-KLRQKAVYIAVLPILSVLSVITHVMSDLINQVVPYCLDNLTMGLAMFLI 273  
 DB 221 LQPHKLKLLVFLSGICVCAIVFLSLSPLEGFALMHTSAVYHIIITMLMNSLWYIN 280  
 QY 274 CEAMSTAHLLAERFQKALKEIGPAAMVADRVLMRLSKLTRDTGNALCTFVMSLYL 333  
 DB 281 SRGIRVASSSLSRCPQOVAIECTAMISRYRFLMNLINSELQALGNVARYSTYCFM 340  
 QY 334 FPIITLSIYGLMSQSG---GFGIDIGLITLALMNIGLFIYCDEAHYASVNRINPCK 390  
 DB 341 FVNITVAIYGLMSQSGFIIDHGFSGFKEIGLYDTVTCSTLIFPCDSNRATIQVAGVQD 400  
 QY 391 KLLMVELMNMNSDAQTEINMFLRATENPSTINGCFEDVNRITLFGKLLTMTTYLVLL 450  
 DB 401 TLISINILKVPQPKKEIDLFQALIEMLPAISLGVAVRELLTSSIAITAIYLVLL 460  
 QY 451 QFOISIPDKDSEGANNTIV 472  
 DB 461 QFKLSISQIIPVEIIEVKKLL 482

RESULT 4  
 Q7PMG3

ID Q7PMG3 PRELIMINARY; PRT; 467 AA.  
 AC Q7PMG3; 01-MAR-2004 (Tremblrel. 26, Created)  
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)  
 DE ENSANGP00000011853.  
 GN Name=ENSANGG00000009364;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 RN NCBI\_TaxID=180454;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -! CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAAB01008980; EMBL13882.2; -  
 SQ SEQUENCE 467 AA; 54183 MW; 20285C3C254E8FEA CRC64;

Query Match 17.3%; Score 458; DB 2; Length 467;  
 Best Local Similarity 30.0%; Pred. No. 7.5e-26;  
 Matches 121; Conservative 82; Mismatches 185; Indels 16; Gaps 9;

QY 65 RLDDYDAPP-LDRKSSDSTASNNPFRKPSVFRNIDPIMFLRIIGVLPYRHGRA-- 121  
 DB 38 QLEQDNESPTMYRKRKIKASDVNLIDQHS-FYHTTKSLVLVQINGVMPIMSPKQVD 96  
 QY 122 --RAKPEMNSAFIYSVFPVLLACVGYANNRHIVRSLSG-PFEERAVLAVFLVNL 178  
 DB 97 MPRITTFWCSAFLMAAFIYACETVILVVARERINPFIISDKRFBVYINIIIMSLMV 156  
 QY 179 PIMIPIIL-WYEARIAKLFDMDFEVLXYOISGHSLPL-KLRQKAVYIAVLPILSVL 236  
 DB 157 PHFLFPVASMNGSEVAKFKNMWTDFOYKYLIVGKPIVPEKLPITWTLCIVMSLSLV 216  
 QY 237 SVITHTVMSDLINQVVPYCLDNLTMGLAM--WFLICEANSTAHLLAERFQKAL 293  
 DB 217 IILSQYLOPDPQFCHTFAYV---HITMLNGFCSLWFVNCTARGTASKARAKLTDVLA 273  
 QY 294 HIGPAAMVADRVLMRLSKLTRDTGNALCTFVMSLYFPIITLSIYGLMSQSG-F 352  
 DB 274 TERPAKLTERRHLMVDSHMQQLGKAYSMTYGIYCVIFPTTIIATYGLSLEIENGA 333  
 QY 353 GIKDIGLITLALMNIGLFIYCDEAHYASVNRINPCKLLMVELMNMNSDAQTEINNF 412  
 DB 334 TYKEVGLFVIFYGMSLFLIICNEAHNASKRGVGNPQRLNVLLTAVDKATQKEVEWFL 393  
 QY 413 RATEENPSTINGCFEDVNRITLFGKLLTMTTYLVLLQFOISI 456  
 DB 394 VAIDKNPPTMNLIDGVANINRGLITSNIFATYLVVLMQFRTL 437

## RESULT 5

G21A\_DROME STANDARD; PRT; 447 AA.  
 ID G21A\_DROME  
 AC G9VPT1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-JAN-2005 (Rel. 46, Last annotation update)  
 DE Putative gustatory receptor 21a.  
 GN Name=G21a; Synonyms=G21D.1; ORNames=G213948;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID=7227;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amentides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Brill J.F., Agbayani A., An H.-J., Nelson C.R., Miklos G.L.G.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,  
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Moberg C., Morris J., Moshiri A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP GENOME REANNOTATION.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminck J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tudy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Beutecourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.B.,  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [3]  
RP IDENTIFICATION AND TISSUE SPECIFICITY  
RX MEDLINE=20175760; PubMed=10710312; DOI=10.1126/science.287.5459.1830;  
RA Clyne P.J., Warr C.G., Carlson J.R.,  
RT "Candidate taste receptors in Drosophila.";  
RL Science 287:1830-1834(2000).  
RN [4]  
RP IDENTIFICATION.  
RX MEDLINE=21407712; PubMed=11516643; DOI=10.1016/S0960-9622(01)00258-5;  
RA Dundupace L., Weisler S., McNealy C., Amrein H.,  
RT "Spatially restricted expression of candidate taste receptors in the  
RT Drosophila gustatory system.";  
RL Curr. Biol. 11:822-835(2001).  
CC -1- FUNCTION: Probable role in the gustatory response.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- TISSUE SPECIFICITY: Expressed in the adult labellar chemosensory  
CC neurons.  
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor Dr-1r  
CC family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF003588; AAF51461.2; -  
DR FBase; FBgn0041250; Gr1a.  
DR GO; GO:0016021; C: integral to membrane; NAS.  
DR GO; GO:0008527; F: taste receptor activity; NAS.  
DR GO; GO:0003009; F: perception of taste; NAS.  
KW G-protein coupled receptor; Glycoprotein; Hypothetical protein;  
KW Multigene family; Receptor; Transmembrane.  
FT DOMAIN 1 107  
FT TRANSMEM 108 128  
FT DOMAIN 129 146  
FT TRANSMEM 147 167  
FT TRANSMEM 168 199  
FT TRANSMEM 200 220  
FT TRANSMEM 221 230  
FT TRANSMEM 231 251  
FT TRANSMEM 252 305  
FT TRANSMEM 306 326  
FT TRANSMEM 327 338  
FT TRANSMEM 339 359  
FT TRANSMEM 360 415  
FT TRANSMEM 416 436  
FT TRANSMEM 437 447  
FT TRANSMEM 448 478  
FT TRANSMEM 479 519  
FT TRANSMEM 520 540  
FT TRANSMEM 541 561  
FT TRANSMEM 562 582  
FT TRANSMEM 583 603  
FT TRANSMEM 604 624  
FT TRANSMEM 625 645  
FT TRANSMEM 646 666  
FT TRANSMEM 667 687  
FT TRANSMEM 688 708  
FT TRANSMEM 709 729  
FT TRANSMEM 730 750  
FT TRANSMEM 751 771  
FT TRANSMEM 772 792  
FT TRANSMEM 793 813  
FT TRANSMEM 814 834  
FT TRANSMEM 835 855  
FT TRANSMEM 856 876  
FT TRANSMEM 877 897  
FT TRANSMEM 898 918  
FT TRANSMEM 919 939  
FT TRANSMEM 940 960  
FT TRANSMEM 961 981  
FT TRANSMEM 982 1002  
FT TRANSMEM 1003 1023  
FT TRANSMEM 1024 1044  
FT TRANSMEM 1045 1065  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1087 1107  
FT TRANSMEM 1108 1128  
FT TRANSMEM 1129 1149  
FT TRANSMEM 1150 1170  
FT TRANSMEM 1171 1191  
FT TRANSMEM 1192 1212  
FT TRANSMEM 1213 1233  
FT TRANSMEM 1234 1254  
FT TRANSMEM 1255 1275  
FT TRANSMEM 1276 1296  
FT TRANSMEM 1297 1317  
FT TRANSMEM 1318 1338  
FT TRANSMEM 1339 1359  
FT TRANSMEM 1360 1380  
FT TRANSMEM 1381 1401  
FT TRANSMEM 1402 1422  
FT TRANSMEM 1423 1443  
FT TRANSMEM 1444 1464  
FT TRANSMEM 1465 1485  
FT TRANSMEM 1486 1506  
FT TRANSMEM 1507 1527  
FT TRANSMEM 1528 1548  
FT TRANSMEM 1549 1569  
FT TRANSMEM 1570 1590  
FT TRANSMEM 1591 1610  
FT TRANSMEM 1611 1631  
FT TRANSMEM 1632 1652  
FT TRANSMEM 1653 1673  
FT TRANSMEM 1674 1693  
FT TRANSMEM 1694 1714  
FT TRANSMEM 1715 1735  
FT TRANSMEM 1736 1756  
FT TRANSMEM 1757 1777  
FT TRANSMEM 1778 1798  
FT TRANSMEM 1799 1818  
FT TRANSMEM 1819 1839  
FT TRANSMEM 1840 1860  
FT TRANSMEM 1861 1881  
FT TRANSMEM 1882 1902  
FT TRANSMEM 1903 1922  
FT TRANSMEM 1923 1943  
FT TRANSMEM 1944 1964  
FT TRANSMEM 1965 1985  
FT TRANSMEM 1986 2006  
FT TRANSMEM 2007 2027  
FT TRANSMEM 2028 2048  
FT TRANSMEM 2049 2069  
FT TRANSMEM 2070 2090  
FT TRANSMEM 2091 2110  
FT TRANSMEM 2111 2131  
FT TRANSMEM 2132 2152  
FT TRANSMEM 2153 2173  
FT TRANSMEM 2174 2194  
FT TRANSMEM 2195 2215  
FT TRANSMEM 2216 2236  
FT TRANSMEM 2237 2257  
FT TRANSMEM 2258 2278  
FT TRANSMEM 2279 2299  
FT TRANSMEM 2300 2320  
FT TRANSMEM 2321 2341  
FT TRANSMEM 2342 2362  
FT TRANSMEM 2363 2383  
FT TRANSMEM 2384 2404  
FT TRANSMEM 2405 2425  
FT TRANSMEM 2426 2446  
FT TRANSMEM 2447 2467  
FT TRANSMEM 2468 2488  
FT TRANSMEM 2489 2509  
FT TRANSMEM 2510 2530  
FT TRANSMEM 2531 2551  
FT TRANSMEM 2552 2572  
FT TRANSMEM 2573 2593  
FT TRANSMEM 2594 2614  
FT TRANSMEM 2615 2635  
FT TRANSMEM 2636 2656  
FT TRANSMEM 2657 2677  
FT TRANSMEM 2678 2698  
FT TRANSMEM 2699 2719  
FT TRANSMEM 2720 2740  
FT TRANSMEM 2741 2761  
FT TRANSMEM 2762 2782  
FT TRANSMEM 2783 2803  
FT TRANSMEM 2804 2824  
FT TRANSMEM 2825 2845  
FT TRANSMEM 2846 2866  
FT TRANSMEM 2867 2887  
FT TRANSMEM 2888 2908  
FT TRANSMEM 2909 2929  
FT TRANSMEM 2930 2950  
FT TRANSMEM 2951 2970  
FT TRANSMEM 2971 2991  
FT TRANSMEM 2992 3012  
FT TRANSMEM 3013 3033  
FT TRANSMEM 3034 3054  
FT TRANSMEM 3055 3075  
FT TRANSMEM 3076 3095  
FT TRANSMEM 3096 3116  
FT TRANSMEM 3117 3137  
FT TRANSMEM 3138 3158  
FT TRANSMEM 3159 3179  
FT TRANSMEM 3180 3200  
FT TRANSMEM 3201 3221  
FT TRANSMEM 3222 3242  
FT TRANSMEM 3243 3263  
FT TRANSMEM 3264 3284  
FT TRANSMEM 3285 3305  
FT TRANSMEM 3306 3326  
FT TRANSMEM 3327 3347  
FT TRANSMEM 3348 3368  
FT TRANSMEM 3369 3389  
FT TRANSMEM 3390 3410  
FT TRANSMEM 3411 3431  
FT TRANSMEM 3432 3452  
FT TRANSMEM 3453 3473  
FT TRANSMEM 3474 3494  
FT TRANSMEM 3495 3515  
FT TRANSMEM 3516 3536  
FT TRANSMEM 3537 3557  
FT TRANSMEM 3558 3578  
FT TRANSMEM 3579 3599  
FT TRANSMEM 3600 3620  
FT TRANSMEM 3621 3641  
FT TRANSMEM 3642 3662  
FT TRANSMEM 3663 3683  
FT TRANSMEM 3684 3704  
FT TRANSMEM 3705 3725  
FT TRANSMEM 3726 3746  
FT TRANSMEM 3747 3767  
FT TRANSMEM 3768 3788  
FT TRANSMEM 3789 3809  
FT TRANSMEM 3810 3830  
FT TRANSMEM 3831 3851  
FT TRANSMEM 3852 3872  
FT TRANSMEM 3873 3893  
FT TRANSMEM 3894 3914  
FT TRANSMEM 3915 3935  
FT TRANSMEM 3936 3956  
FT TRANSMEM 3957 3977  
FT TRANSMEM 3978 3998  
FT TRANSMEM 3999 4019  
FT TRANSMEM 4020 4040  
FT TRANSMEM 4041 4061  
FT TRANSMEM 4062 4082  
FT TRANSMEM 4083 4103  
FT TRANSMEM 4104 4124  
FT TRANSMEM 4125 4145  
FT TRANSMEM 4146 4166  
FT TRANSMEM 4167 4187  
FT TRANSMEM 4188 4208  
FT TRANSMEM 4209 4229  
FT TRANSMEM 4230 4250  
FT TRANSMEM 4251 4270  
FT TRANSMEM 4271 4291  
FT TRANSMEM 4292 4312  
FT TRANSMEM 4313 4333  
FT TRANSMEM 4334 4354  
FT TRANSMEM 4355 4375  
FT TRANSMEM 4376 4396  
FT TRANSMEM 4397 4417  
FT TRANSMEM 4418 4438  
FT TRANSMEM 4439 4459  
FT TRANSMEM 4460 4480  
FT TRANSMEM 4481 4501  
FT TRANSMEM 4502 4522  
FT TRANSMEM 4523 4543  
FT TRANSMEM 4544 4564  
FT TRANSMEM 4565 4585  
FT TRANSMEM 4586 4606  
FT TRANSMEM 4607 4627  
FT TRANSMEM 4628 4648  
FT TRANSMEM 4649 4669  
FT TRANSMEM 4670 4690  
FT TRANSMEM 4691 4711  
FT TRANSMEM 4712 4732  
FT TRANSMEM 4733 4753  
FT TRANSMEM 4754 4774  
FT TRANSMEM 4775 4795  
FT TRANSMEM 4796 4816  
FT TRANSMEM 4817 4837  
FT TRANSMEM 4838 4858  
FT TRANSMEM 4859 4879  
FT TRANSMEM 4880 4900  
FT TRANSMEM 4901 4921  
FT TRANSMEM 4922 4942  
FT TRANSMEM 4943 4963  
FT TRANSMEM 4964 4984  
FT TRANSMEM 4985 5005  
FT TRANSMEM 5006 5026  
FT TRANSMEM 5027 5047  
FT TRANSMEM 5048 5068  
FT TRANSMEM 5069 5089  
FT TRANSMEM 5090 5110  
FT TRANSMEM 5111 5131  
FT TRANSMEM 5132 5152  
FT TRANSMEM 5153 5173  
FT TRANSMEM 5174 5194  
FT TRANSMEM 5195 5215  
FT TRANSMEM 5216 5236  
FT TRANSMEM 5237 5257  
FT TRANSMEM 5258 5278  
FT TRANSMEM 5279 5299  
FT TRANSMEM 5300 5320  
FT TRANSMEM 5321 5341  
FT TRANSMEM 5342 5362  
FT TRANSMEM 5363 5383  
FT TRANSMEM 5384 5404  
FT TRANSMEM 5405 5425  
FT TRANSMEM 5426 5446  
FT TRANSMEM 5447 5467  
FT TRANSMEM 5468 5488  
FT TRANSMEM 5489 5509  
FT TRANSMEM 5510 5530  
FT TRANSMEM 5531 5551  
FT TRANSMEM 5552 5572  
FT TRANSMEM 5573 5593  
FT TRANSMEM 5594 5614  
FT TRANSMEM 5615 5635  
FT TRANSMEM 5636 5656  
FT TRANSMEM 5657 5677  
FT TRANSMEM 5678 5698  
FT TRANSMEM 5699 5719  
FT TRANSMEM 5720 5740  
FT TRANSMEM 5741 5761  
FT TRANSMEM 5762 5782  
FT TRANSMEM 5783 5803  
FT TRANSMEM 5804 5824  
FT TRANSMEM 5825 5845  
FT TRANSMEM 5846 5866  
FT TRANSMEM 5867 5887  
FT TRANSMEM 5888 5908  
FT TRANSMEM 5909 5929  
FT TRANSMEM 5930 5950  
FT TRANSMEM 5951 5970  
FT TRANSMEM 5971 5991  
FT TRANSMEM 5992 6012  
FT TRANSMEM 6013 6033  
FT TRANSMEM 6034 6054  
FT TRANSMEM 6055 6075  
FT TRANSMEM 6076 6096  
FT TRANSMEM 6097 6117  
FT TRANSMEM 6118 6138  
FT TRANSMEM 6139 6159  
FT TRANSMEM 6160 6180  
FT TRANSMEM 6181 6201  
FT TRANSMEM 6202 6222  
FT TRANSMEM 6223 6243  
FT TRANSMEM 6244 6264  
FT TRANSMEM 6265 6285  
FT TRANSMEM 6286 6306  
FT TRANSMEM 6307 6327  
FT TRANSMEM 6328 6348  
FT TRANSMEM 6349 6369  
FT TRANSMEM 6370 6390  
FT TRANSMEM 6391 6411  
FT TRANSMEM 6412 6432  
FT TRANSMEM 6433 6453  
FT TRANSMEM 6454 6474  
FT TRANSMEM 6475 6495  
FT TRANSMEM 6496 6516  
FT TRANSMEM 6517 6537  
FT TRANSMEM 6538 6558  
FT TRANSMEM 6559 6579  
FT TRANSMEM 6580 6600  
FT TRANSMEM 6601 6621  
FT TRANSMEM 6622 6642  
FT TRANSMEM 6643 6663  
FT TRANSMEM 6664 6684  
FT TRANSMEM 6685 6705  
FT TRANSMEM 6706 6726  
FT TRANSMEM 6727 6747  
FT TRANSMEM 6748 6768  
FT TRANSMEM 6769 6789  
FT TRANSMEM 6790 6810  
FT TRANSMEM 6811 6831  
FT TRANSMEM 6832 6852  
FT TRANSMEM 6853 6873  
FT TRANSMEM 6874 6894  
FT TRANSMEM 6895 6915  
FT TRANSMEM 6916 6936  
FT TRANSMEM 6937 6957  
FT TRANSMEM 6958 6978  
FT TRANSMEM 6979 6999  
FT TRANSMEM 7000 7020  
FT TRANSMEM 7021 7041  
FT TRANSMEM 7042 7062  
FT TRANSMEM 7063 7083  
FT TRANSMEM 7084 7104  
FT TRANSMEM 7105 7125  
FT TRANSMEM 7126 7146  
FT TRANSMEM 7147 7167  
FT TRANSMEM 7168 7188  
FT TRANSMEM 7189 7209  
FT TRANSMEM 7210 7230  
FT TRANSMEM 7231 7251  
FT TRANSMEM 7252 7272  
FT TRANSMEM 7273 7293  
FT TRANSMEM 7294 7314  
FT TRANSMEM 7315 7335  
FT TRANSMEM 7336 7356  
FT TRANSMEM 7357 7377  
FT TRANSMEM 7378 7398  
FT TRANSMEM 7399 7419  
FT TRANSMEM 7420 7440  
FT TRANSMEM 7441 7461  
FT TRANSMEM 7462 7482  
FT TRANSMEM 7483 7503  
FT TRANSMEM 7504 7524  
FT TRANSMEM 7525 7545  
FT TRANSMEM 7546 7566  
FT TRANSMEM 7567 7587  
FT TRANSMEM 7588 7608  
FT TRANSMEM 7609 7629  
FT TRANSMEM 7630 7650  
FT TRANSMEM 7651 7671  
FT TRANSMEM 7672 7692  
FT TRANSMEM 7693 7713  
FT TRANSMEM 7714 7734  
FT TRANSMEM 7735 7755  
FT TRANSMEM 7756 7776  
FT TRANSMEM 7777 7797  
FT TRANSMEM 7798 7818  
FT TRANSMEM 7819 7839  
FT TRANSMEM 7840 7860  
FT TRANSMEM 7861 7881  
FT TRANSMEM 7882 7902  
FT TRANSMEM 7903 7923  
FT TRANSMEM 7924 7944  
FT TRANSMEM 7945 7965  
FT TRANSMEM 7966 7986  
FT TRANSMEM 7987 8007  
FT TRANSMEM 8008 8028  
FT TRANSMEM 8029 8049  
FT TRANSMEM 8050 8070  
FT TRANSMEM 8071 8091  
FT TRANSMEM 8092 8112  
FT TRANSMEM 8113 8133  
FT TRANSMEM 8134 8154  
FT TRANSMEM 8155 8175  
FT TRANSMEM 8176 8196  
FT TRANSMEM 8197 8217  
FT TRANSMEM 8218 8238  
FT TRANSMEM 8239 8259  
FT TRANSMEM 8260 8280  
FT TRANSMEM 8281 8301  
FT TRANSMEM 8302 8322  
FT TRANSMEM 8323 8343  
FT TRANSMEM 8344 8364  
FT TRANSMEM 8365 8385  
FT TRANSMEM 8386 8406  
FT TRANSMEM 8407 8427  
FT TRANSMEM 8428 8448  
FT TRANSMEM 8449 8469  
FT TRANSMEM 8470 8490  
FT TRANSMEM 8491 8511  
FT TRANSMEM 8512 8532  
FT TRANSMEM 8533 8553  
FT TRANSMEM 8554 8574  
FT TRANSMEM 8575 8595  
FT TRANSMEM 8596 8616  
FT TRANSMEM 8617 8637  
FT TRANSMEM 8638 8658  
FT TRANSMEM 8659 8679  
FT TRANSMEM 8680 8700  
FT TRANSMEM 8701 8721  
FT TRANSMEM 8722 8742  
FT TRANSMEM 8743 8763  
FT TRANSMEM 8764 8784  
FT TRANSMEM 8785 8805  
FT TRANSMEM 8806 8826  
FT TRANSMEM 8827 8847  
FT TRANSMEM 8848 8868  
FT TRANSMEM 8869 8889  
FT TRANSMEM 8890 8910  
FT TRANSMEM 8911 8931  
FT TRANSMEM 8932 8952  
FT TRANSMEM 8953 8973  
FT TRANSMEM 8974 8994  
FT TRANSMEM 8995 9015  
FT TRANSMEM 9016 9036  
FT TRANSMEM 9037 9057  
FT TRANSMEM 9058 9078  
FT TRANSMEM 9079 9099  
FT TRANSMEM 9100 9120  
FT TRANSMEM 9121 9141  
FT TRANSMEM 9142 9162  
FT TRANSMEM 9163 9183  
FT TRANSMEM 9184 9204  
FT TRANSMEM 9205 9225  
FT TRANSMEM 9226 9246  
FT TRANSMEM 9247 9267  
FT TRANSMEM 9268 9288  
FT TRANSMEM 9289 9309  
FT TRANSMEM 9310 9330  
FT TRANSMEM 9331 9351  
FT TRANSMEM 9352 9372  
FT TRANSMEM 9373 9393  
FT TRANSMEM 9394 9414  
FT TRANSMEM 9415 9435  
FT TRANSMEM 9436 9456  
FT TRANSMEM 9457 9477  
FT TRANSMEM 9478 9498  
FT TRANSMEM 9499 9519  
FT TRANSMEM 9520 9540  
FT TRANSMEM 9541 9561  
FT TRANSMEM 9562 9582  
FT TRANSMEM 9583 9603  
FT TRANSMEM 9604 9624  
FT TRANSMEM 9625 9645  
FT TRANSMEM 9646 9666  
FT TRANSMEM 9667 9687  
FT TRANSMEM 9688 9708  
FT TRANSMEM 9709 9729  
FT TRANSMEM 9730 9750  
FT TRANSMEM 9751 9771  
FT TRANSMEM 9772 9792  
FT TRANSMEM 9793 9813  
FT TRANSMEM 9814 9834  
FT TRANSMEM 9835 9855  
FT TRANSMEM 9856 9876  
FT TRANSMEM 9877 9897  
FT TRANSMEM 9898 9918  
FT TRANSMEM 9919 9939  
FT TRANSMEM 9940 9960  
FT TRANSMEM 9961 9981  
FT TRANSMEM 9982 10000  
FT TRANSMEM 10001 10021  
FT TRANSMEM 10022 10042  
FT TRANSMEM 10043 10063  
FT TRANSMEM 10064 10084  
FT TRANSMEM 10085 10105  
FT TRANSMEM 10106 10126  
FT TRANSMEM 10127 10147  
FT TRANSMEM 10148 10168  
FT TRANSMEM 10169 10189  
FT TRANSMEM 10190 10210  
FT TRANSMEM 10211 10231  
FT TRANSMEM 10232 10252  
FT TRANSMEM 10253 10273  
FT TRANSMEM 10274 10294  
FT TRANSMEM 10295 10315  
FT TRANSMEM 10316 10336  
FT TRANSMEM 10337 10357  
FT TRANSMEM 10358 10378  
FT TRANSMEM 10379 10399  
FT TRANSMEM 10400 10420  
FT TRANSMEM 10421 10441  
FT TRANSMEM 10442 10462  
FT TRANSMEM 10463 10483  
FT TRANSMEM 10484 10504  
FT TRANSMEM 10505 10525  
FT TRANSMEM 10526 10546  
FT TRANSMEM 10547 10567  
FT TRANSMEM 10568 10588  
FT TRANSMEM 10589 10609  
FT TRANSMEM 10610 10630  
FT TRANSMEM 10631 10651  
FT TRANSMEM 10652 10672  
FT TRANSMEM 10673 10693  
FT TRANSMEM 10694 10714  
FT TRANSMEM 10715 10735  
FT TRANSMEM 10736 10756  
FT TRANSMEM 10757 10777  
FT TRANSMEM 10778 10798  
FT TRANSMEM 10799 10819  
FT TRANSMEM 10820 10840  
FT TRANSMEM 10841 10861  
FT TRANSMEM 10862 10882  
FT TRANSMEM 10883 10903  
FT TRANSMEM 10904 10924  
FT TRANSMEM 10925 10945  
FT TRANSMEM 10946 10966  
FT TRANSMEM 10967 10987  
FT TRANSMEM 10988 11008  
FT TRANSMEM 11009 11029  
FT TRANSMEM 11030 11050  
FT TRANSMEM 11051 11071  
FT TRANSMEM 11072 11092  
FT TRANSMEM 11093 11113  
FT TRANSMEM 11114 11134  
FT TRANSMEM 11135 11155  
FT TRANSMEM 11156 11176  
FT TRANSMEM 11177 11197  
FT TRANSMEM 11198 11218  
FT TRANSMEM 11219 11239  
FT TRANSMEM 11240 11260  
FT TRANSMEM 11261 11281  
FT TRANSMEM 11282 11302  
FT TRANSMEM 11303 11323  
FT TRANSMEM 11324 11344  
FT TRANSMEM 11345 11365  
FT TRANSMEM 11366 11386  
FT TRANSMEM 11387 11407  
FT TRANSMEM 11408 11428  
FT TRANSMEM 11429 11449  
FT TRANSMEM 11450 11470  
FT TRANSMEM 11471 11491  
FT TRANSMEM 11492 11512  
FT TRANSMEM 11513 11533  
FT TRANSMEM 11534 11554  
FT TRANSMEM 11555 11575  
FT TRANSMEM 11576 11596  
FT TRANSMEM 11597 11617  
FT TRANSMEM 11618 11638  
FT TRANSMEM 11639 11659  
FT TRANSMEM 11660 11680  
FT TRANSMEM 11681 11701  
FT TRANSMEM 11702 11722  
FT TRANSMEM 11723 11743  
FT TRANSMEM 11744 11764  
FT TRANSMEM 11765 11785  
FT TRANSMEM 11786 11806  
FT TRANSMEM 11807 11827  
FT TRANSMEM 11828 11848  
FT TRANSMEM 11849 11869  
FT TRANSMEM 11870 11890  
FT TRANSMEM 11891 11911  
FT TRANSMEM 11912 11932  
FT TRANSMEM 11933 11953  
FT TRANSMEM 11954 11974  
FT TRANSMEM 11975 11995  
FT TRANSMEM 11996 12016  
FT TRANSMEM 12017 12037  
FT TRANSMEM 12038 12058  
FT TRANSMEM 12059 12079  
FT TRANSMEM 12080 12100  
FT TRANSMEM 12101 12121  
FT TRANSMEM 12122 12142  
FT TRANSMEM 12143 12163  
FT TRANSMEM 12164 12184  
FT TRANSMEM 12185 12205  
FT TRANSMEM 12206 12226  
FT TRANSMEM 12227 12247  
FT TRANSMEM 12248 12268  
FT TRANSMEM 12269 12289  
FT TRANSMEM 12290 12310  
FT TRANSMEM 12311 12331  
FT TRANSMEM 12332 12352  
FT TRANSMEM 12353 12373  
FT TRANSMEM 12374 12394  
FT TRANSMEM 12395 12415  
FT TRANSMEM 12416 12436  
FT TRANSMEM 12437 12457  
FT TRANSMEM 12458 12478  
FT TRANSMEM 12479 12499  
FT TRANSMEM 12500 12520  
FT TRANSMEM 12521 12541  
FT TRANSMEM 12542 12562  
FT TRANSMEM 12563 12583  
FT TRANSMEM 12584 12604  
FT TRANSMEM 12605 12625  
FT TRANSMEM 12626 12646  
FT TRANSMEM 12647 12667  
FT TRANSMEM 12668 12688  
FT TRANSMEM 12689 12709  
FT TRANSMEM 12710 12730  
FT TRANSMEM 12731 12751  
FT TRANSMEM 12752 12772  
FT TRANSMEM 12773 12793  
FT TRANSMEM 12794 12814  
FT TRANSMEM 12815 12835  
FT TRANSMEM 12836 12856  
FT TRANSMEM 12857 12877  
FT TRANSMEM 12878 12898  
FT TRANSMEM 12899 12919  
FT TRANSMEM 12920 12940  
FT TRANSMEM 12941 12961  
FT TRANSMEM 12962 12982  
FT TRANSMEM 12983 13003  
FT TRANSMEM 13004 13024  
FT TRANSMEM 13025 13045  
FT TRANSMEM 13046 13066  
FT TRANSMEM 13067 13087  
FT TRANSMEM 13088 13108  
FT TRANSMEM 13109 13129  
FT TRANSMEM 13130 13150  
FT TRANSMEM 13151 13171  
FT TRANSMEM 13172 13192  
FT TRANSMEM 13193 13213  
FT TRANSMEM 13214 13234  
FT TRANSMEM 13235 13255  
FT TRANSMEM 13256 13276  
FT TRANSMEM 13277 132

DT 05-JUN-2004 (Rel. 44, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Putative gustatory receptor 98d.  
GN Name=Gr98d; Synonyms=GR98B.1; ORFNames=CG31061;  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
RN RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley.  
RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
EX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanaidee P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Baau A.A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris B.A., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jatalai M., Kalish P., Karen G.H., Ke Z., Kension J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Milshina N.V., Mobery C., Morris J., Mostrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spraker B., Spradling A.C., Stepleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S., Woodedge T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
RN [2]  
RP GENOME REANNOTATION.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kankar J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stables M., Yamada C., Ashburner W., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE=20175760; PubMed=10710312; DOI=10.1126/science.287.5459.1830;  
RA Clyne P.J., Warr C.G., Carlson J.R.;  
RT "Candidate taste receptors in *Drosophila*.";  
RL Science 287:1830-1834(2000).  
CC -1- FUNCTION: Probable role in the gustatory response.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor Dr-tr

[illegible]



```

Db 39 LGVPIYERH-PATL-----SLQFRSLLILCYFIALIGPMRCMVLTTFPTFNFGVSA 92
Qy 170 AY---LPLVNTLPLMIIP-LIMTEARKAKLPNDMDPEVLYIQISGSLPLKAKQAV 224
Db 93 HYTAMFMEILDAILLVPPCYLLCRHKKAKELFEVLV--VKVYSSLSLSLFPRLMFRAC 150
Qy 225 YIAYLPLISLVSVITHTVMSDMLNNOVVP-----YCLDNLTLTML----- 266
Db 151 FL-----FNLVYDSAYGVNVALNLVLPETGHEKYVKYLISWSVLSQLTISIYLLI 201
Qy 267 -GAMFLLICEAMSIYHLLAER-----FOKALK-HISPAAMVADRYVLMLSLKLTED 317
Db 202 YGTAYQLRLMDHDLMDHLAAGRLSTMYVQEOIRQGHGPRAGVLEFYBQUG-KMSFTLND 260
Qy 318 TGNALCYTFVFNSL-YLFFIITLSIYGLMSQUS---EGFGKIDGLTTLTALNIGLLEY 372
Db 261 LYGVDPLITFLMNTLHLHLEFVCYMTLTKVLTKNSYLSMHALLLIGIAVTSVNDLLCLMKI 320
Qy 373 ICDEAHYASVNVRTNFOKGLMLVVELNMWMSDAOTEINMFLPARENNPSTINCQGFVDNR 432
Db 321 VGTFRARTREESIKT--QKLLRLINTLSPMDHKQSIIEVALQTLHOPIEFTACRMPTLDY 378
Qy 433 TLFGKLLTMVTYVLVLTLOFOISI 456
Db 379 TVLFSIAAAVTYLLIILIOFEMAI 402

```

<b>RESULT 10</b>					
ID	O7PPB5	PRELIMINARY;	PRT;	444 AA.	
AC	O7PPB5:				
DT	01-MAR-2004 (TREMBLrel. 26, Created)				
DJ	01-MAR-2004 (TREMBLrel. 26, Last sequence update)				
DR	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	ENSANGRO0000004263.				
CN	Name=ENSANGRO000000364;				
OS	Anopheles gambiae str. PEST.				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
CC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
OX	NCBI_TaxID=180454;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=PEST;				
FA	Anopheles Genome Sequencing Consortium;				
RL	Submitted (APR-2003) to the EMBL/genbank/DBJ databases.				
CC	-!- CAUTION: The sequence shown here is derived from an				
CC	EMBL/genbank/DBJ whole genome shotgun (WGS) entry which is				
CC	preliminary data.				
DR	EMBL; AAA01008960; EAA10739.2; -.				
DR	InterPro; IPR009318; Trehalose_recpt.				
DR	Pfam; PF06151; Trehalose_recpt; 1.				
SQ	SEQUENCE 444 AA: 51486 MW; 6FE09ACD0EE059A9 CRC64;				
	Query Match	5.4%; Score 143; DB 2; Length 444;			
	Best Local Similarity	21.5%; Pred. No. 0.023;			
	Matches	98; Conservative 73; Mismatches 170; Indels 114; Gaps			
OY	59 SIGLAERLDADYEAPPLD---RKKSDDSYASNNEPFKPSVRYRNIDPIMFLRIIGVLPI	115			
		:	:	:	:
DB	26 STALAE-----ETPPADGAARECGSTHEAVAIWMQLF--SLIPDIGVAR-----	70			
OY	116 VHHGPARAKFNENNSAFYS-VVFPLLACYGYVANNR.IHIVRSISGFPERA--VIAYL	172			
DB	71 -STDPDVMRRISRVOFYGVCTLFMLTLITLMC---VHTNAEFSFGVQNTSLVYA	125			
OY	173 PLV-NILPMIIPILMYEARAKTAKFNMDDEEVLYIQISGLSPKDKOKAYIAIVLP	231			
	: : : : :	:	:	:	:
DB	126 IIVFFNVELMLLARWN-SQINGRWYTDEAPFRTPDYRPRSRLF--PRKKVHLIAFGVM	181			
OY	232 ILISLVSVITHVTMSLN-----INQVPVICIDNTLMLG	267			
	: : : : :	:	:	:	:
DB	182 FLAFEDDTLNFSAVRILDELIRYCPTAGFWKNFPHREHPYLARIYPHVVGWGTIELT	241			

```

OY 268 -----AW-----PFCAMSIITALLBLERPOKAKHI  GPMAMADYRVL-----WLRSK 313
Db 242 MRAIKETWTHYVDVFIFICLSLQHRFFV--QFNERLERIDCGPQSGCVMRALRIDYVRLSE 299
OY 314 LTRDTGNALCYTFV--FMSLYLF-----PITTSIYGLMSQLSBGPICIKOIGLTITVL 364
Db 300 LV-----TFVDERFSKLILPFCANDMFPITV-----OLFNSFLKPTITVITYVF 343
OY 365 W-NTGLFPYICDEAHYASVNVRTNPOKKLAW-----ELUMWMSDAQTEINMFJR-----AT 415
Db 344 WYSLGFLILGRCPGLMLFVVSSISRASEKULETTRRPPSTW-----NLDLRRLCDAY 394
OY 416 EMNPSTINGCFPDVNRTLFKGLITTMVYYVLVL 450
Db 395 ATSEWALSGKRFVVRPLILAMAGTIIITYBELVL 429

```

ID	07PK14	PRELIMINARY;	PRT;	416 AA.
AC	07PK14			
DT	01-MAR-2004 (TREMBLrel. 26, Created)			
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	ENSANGP0000023329.			
CN	Name=ENSANGS000020045;			
OS	Anopheles gambiae str. PEST.			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.			
OX	NCBI_TaxID=180454;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PEST;			
RA	Anopheles Genome Sequencing Consortium;			
RL	Submitted (Apr-2003) to the EMBL/Genbank/DBJ databases.			
CC	-!- CAUTION: The sequence shown here is derived from an			
CC	EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is			
CC	preliminary data.			
DR	EMBL; AAA01008987; EAA43247.1; -			
SQ	SEQUENCE 416 AA; 48071 MW; 1A6030890A0F68B1 CRC64;			
Query Match	5.4%; Score 142; DB 2; Length 416;			
Best Local Similarity	19.6%; Pred. No. 0.025;			
Matches	81; Conservative 80; Mismatches 182; Indels 70; Gaps 15;			
QY	100 IDPNNWLR-----IIGVLEPIV--RHGPAPAKFEMNSASPLY 134			
DB	12 VSPFMSVRRRPSCCTMOVPRMLIVLQSVLKLFLALPLSPDR--RRFRNRKYDVL 68			
QY	135 SVFPEVLLACVGYVANNRIHVR-SLSGPFEEAVIATLFLVNIIPIIMIIPTAEARK- 192			
DB	69 STVLVLVLA-IVGPVVCVELVYMNISVENQITNLSAQPAFVFFIILVOIIFIRKKAC 127			
QY	193 IAKLFNDWDPEVLYYQISGSLPKIRQKAVYIAIVLPILSVLSVITHTMSDLNINQ 252			
DB	128 LHVLTNMFQKHVLRVAVARSM-LEFRYLRLFLVKYAVV-DILMLALATYFEDSTYTG 185			
QY	253 VVPYCIIDNLNRMGLAMWFLICEAMSTIAHLIAERFQALKHIGPAA----- 300			
DB	186 RVDVSGLSMWLTFFYAMFFIAVENFIAGVILGAVQMAVNTIGKRLARSRQSAAGRQ 245			
QY	301 --VADRYVLMRLSLKTRDGTGNAICYTFVFNLSYLFPIITISYIGMSQLSGFGIKDIG 358			
DB	246 PSVLQVYLKHCNEMVYQFMEITLNPFLTMLTGWTFMFIYVSVRPASPV--GNGIAD-- 301			
QY	359 LTTTALMIGLIF-YICDEAHYASV--NVRTNFQKKLMLVELMNKNSDAQIE----- 407			
DB	302 -DFKAYINPLIFELQCVQLYLVLIIPSVYTDHAKQWRLT-LNYSVAVQHRRPGQERLV 359			
QY	408 -----IMPLFATEMNPSTINCGFPPVNRRLTFGKLITWTATYVVLVLLQFOI 454			
DB	360 GREGGIVLMDCKQRRNNAIINNYGMVAMNRLLFGMAITMTSYLIIILQFPI 412			

RESULT 12					
ID	Q7PVK0	PRELIMINARY;	PRT;	411 AA.	
AC	Q7PVJ8				
DT	01-MAR-2004 (TREMBLrel. 26, Created)				
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)				
DE	ENSANGP0000023884.				
GN	Name=ENSANGC0000020468;				
OS	Anopheles gambiae scr. PEST.				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
NCBI_TaxID=180454;	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=PEST;				
RA	Anopheles Genome Sequencing Consortium;				
RL	Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.				
CC	-!- CAUTION: The sequence shown here is derived from an				
CC	EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is				
DR	preliminary data.				
SQ	EMBL; AAA501008964; EAA43430.1; --				
	SEQUENCE 411 AA; 47111 MW; 60D7A379CB2729E5E CRC64;				
	Query Match 5.3%; Score 141.5; DB 2; Length 411,				
	Best Local Similarity 21.8%; Pred. No. 0.027;				
	Matches 87; Conservative 64; Mismatches 161; Indels 87; Gaps 16;				
QY	110 IGVLPVHGHGPARAKFENMSAFIYSVFV---LLACTV---GYANNRIHIYRSISG 162				
Dd	46 LGVVPEYEPH-DATLSLOFSRSLVYCFTIAIISPGRCCVFELTPPTWKLGISVVR-YTA 103				
QY	163 PFEEAVLAIVLELVNLPIMLIP--ILWEARKIALPFNDMDPEVLYYQIGSHPL--- 217				
Dd	104 WREHLDAAIL-----LVIPCYLCQHAKVKQLFE-----LSVEIRSSSLPGIF 148				
QY	218 -----KLQKAVY-----IAVLPILSVAIVTHVTMTSDINQNVPCILD 260				
Dd	149 PFMQMFRCLLCSLVDVSVCVNNAFNVLVGAGHNAYV-----QILLMSYL 197				
QY	261 NLT-----AMLGAWMFICRAMSTITALAE-----FQKALK-HIGPAWADYRV 306				
Dd	198 QLTTSVVLIIYGTAQYLLMMWDQKHDLVAGRSLTNVPVOBOIROOHNGPAGYLMEYEQ 257				
QY	307 LMLRLSKLTRTGNAICTFYFMSLYLEFIITLSIGM-----SOLSEGRGKDI 357				
Dd	258 LAKVCEITNDLIQVPLITYFLMALIHILEFCYLLITLKLVMSYTTWOSILIAGAIASY 317				
QY	358 GLATTALNNIGLIFYICDEAHYASVYNVTEQKLLMWELMNNSDAOTEINMETLATTEM 417				
Dd	318 VIDTLCLMRIYGTf-----ARTRESLKT--QKLLIRLNTSPMDHKLQKSIEVFALQTIL 370				
QY	418 NPSTINGCGFPDVNRTEFKGLITTWVLYLVLLQFOISI 456				
Dd	371 QPIETACRMFTLDYTVLFISAADVNTYLLIIOFEMAI 409				
RESULT 13					
ID	Q7PVJ8	PRELIMINARY;	PRT;	382 AA.	
AC	Q7PVJ8				
DT	01-MAR-2004 (TREMBLrel. 26, Created)				
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	ENSANGP0000024480.				
GN	Name=ENSANGC0000020468;				
OS	Anopheles gambiae scr. PEST.				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
NC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
NCBI_TaxID=180454;					
RN	[1]				
RP	SEQUENCE FROM N.A.				

RC	STRAIN=PEST;	Score 140;	DB 2;	Length 382;
RA	Anopheles Genome Sequencing Consortium;			
RL	Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.			
CC	-1- CAUTION: The sequence shown here is derived from an			
CC	EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is			
CC	preliminary data.			
DR	EMBL; AAB01008984; EAA4343.1; -			
DR	SEQUENCE 382 AA; 43445 MW; 8863D7248E834FBC CRC64;			
SQL				
Query Match	5.3%; Score 140; DB 2; Length 382;			
Best Local Similarity	20.4%; Pred. No. 0.032;			
Matches	81; Conservative 78; Mismatches 155; Indels 84; Gaps 18			
OY	109 IIGVLPYVRHGPAPAKFEFEMNSAFYSVVFVLLACTGYGVANNRHHVRSLSGPEEAV 168			
DB	17 LVGVIPL--RLRPSWEL-----LILCYTIALAGPLVHLLVLAM-FYIVM 61			
OY	169 IAY--LFLVNI-----LPIIIPILMY---EARKIAKLFNMDPEVLYYQISGS- 214			
DB	62 VKFRFTFVVSICMHLGEVPLIFPPCYLLSCRARKKOLAELCTVHSSSLVGPESHAM 121			
OY	215 ----LPLKLRKAYIAIVLPIILSVLS-----VITHWMSDINNOVPCYL 259			
DB	122 YRKFCEVNLGYEAHATVTLNLNLSQVFPFGMGKIYQTLMLNLSLTSYAVL----- 176			
OY	260 DNLTPAMLGAMWFLICEAMSIHAHLAERF-----QALKR--HIGPAMVADRYVLM 309			
DB	177 --LTMVGSVQYLMIC---AVKLNLLEPLRLSSWTDPRYGETLRLHFGRDGLMDFYGLM- 230			
OY	310 RLKSLTRDTGALCTFVFMNLYLFFITLSTI-YGLMSQLSBGG---IDIGLITIAL 364			
DB	231 ----QGAEIANELFGAPLALAYIAIAFHTTIIYYRVMSQLAGLENQPLVTHMLIVLV 286			
OY	365 W-NIGLFEYI-----CDEAHYASVNVRTNFGKLLIMELNMNNSDAQEINMFRATEMN 418			
DB	287 WVTMDLWFLMAVVGSCQGMRESLKT-----QKLLRLNLSMDHKXQSLVFPALQTLHQ 342			
OY	419 PSTNCGGFDPVNRFTLFEKGLITVTWYIVLVLLQFOIST 456			
DB	343 PIEFTACRMFTLDYTVLFSIAAATVNYLLILOFEMAI 380			
RESULT 14				
OYPKCS	PRELIMINARY; PRT; 451 AA.			
ID	OYPKCS			
AC	OYPKCS;			
DT	01-MAR-2004 (TREMBLrel. 26, Created)			
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	ENSANGP0000022858			
GN	Name=ENSANGP0000022858			
OS	Anopheles gambiae str. PEST.			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
CC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.			
CC	NCBI_TaxId=180454;			
CC	NCBI_TaxId=180454;			
CC	SEQUENCE FROM N.A.			
CC	STRAIN=PEST;			
RA	Anopheles Genome Sequencing Consortium;			
RL	Submitted (APR-2003) to the EMBL/Genbank/DBJ databases.			
CC	-1- CAUTION: The sequence shown here is derived from an			
CC	EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is			
CC	preliminary data.			
DR	EMBL; AAB01008986; EAA43315.1; -			
DR	SEQUENCE 451 AA; 51724 MW; 4CF7E1BDB849A84 CRC64;			
SQL				
Query Match	5.3%; Score 139.5; DB 2; Length 451;			
Best Local Similarity	18.7%; Pred. No. 0.042;			
Matches	86; Conservative 95; Mismatches 174; Indels 105; Gaps 21			
OY	72 APPIDRKSSDSVTSANNPEKPSVFRNIDINMFLRIIGVLPYVRHGPAPAKFEMNSAS 131			
DB	16 APPKRPGLAQTAVGEOEPE-QLFHFAFKCFRLFALTPLGIM-----DRQDQRYRVNTR 69			

```

QY 132 PIYSVVFVLLACVYG-----YANNRIHVRSLSGFEEAVIAYLVNLIPI---MIIP 184
DB 70 WMLLVLLVAVVAMIALFETFFIERRTALITGIANHIQ-----FLMNTIALTVAMVVF 122
QY 185 ILMYERAKIAKLFNDDEVLVYQISGHS-----LPLRKQAVYIAVLILSV 235
DB 123 QL-----KADELGSILIDGFLILIDRELSYVHVEVAGYKRVSFLLRGVALLAM--SLTV 175
QY 236 LSAVITHTMSDINI-----NOVVPYCI--LDNITAMIGAMWF-----LICEAMSTTA 281
DB 176 YGDFVSFVQITVYEVVYMLSHQPLIITYAMAFIAVILYIMHARRRLNTLVEQYVRQG 235
QY 282 HLLAERFQ-----KALKHIGPAMVADYRVLMRLSLKLTPTDGNALC 323
DB 236 HIFAPARQITISPATWKLDEBSAVGBELHSVGRREISDPLQV--LAIVSRITDLGQKIE 293
QY 324 YTF--VFMSLY--LFIITLTIY-----GLMSQISEGFGIKDI---GLTITLMTNIGLTF 371
DB 294 SYFGPLFVTVYTLFVTVTIQSYCYALHTAKDREGISTELVSGGIIILYNVIAIYALP 353
QY 372 YICDEAHYASVNVRTNFOKKLTM---VELNMNNSDAQTEINMFLRATENMNSTI-----N 423
DB 354 YICEQVSEBS-----KILMSYLSKLSMKHSQVAGHSIWP-----PNLISSVRFPS 398
QY 424 CGGFEDVNRITLFXGLITTVTVYLVLLQFOISIPDKGDS 463.
DB 399 AFGFPTINYMMLSGLVAGVTVYLIIFIQFNSWVPAGKDDT 438

```

## RESULT 15

```

O7OKLO PRELIMINARY; PRT; 1210 AA.
AC 07OKLO;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Eblp4099 (Fragment).
GN Name=eblg4099; ORFNames=ENSANG0000003253;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prexygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_taxid=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RU Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008794; EAA03574.1; -.
DR InterPro: IPR009318; Trehalose_recept.
DR Pfam: PF06151; Trehalose_recp. 2.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 1210 AA; 140580 MW; 6DF7CA521861C008 CRC64;

```

```

Query Match 5.3%; Score 139.5; DB 2; Length 1210;
Best Local Similarity 22.7%; Pred. No. 0.13; Indels 117; Gaps 20;
Matches 98; Conservative 62; Mismatches 154;

```

```

QY 108 RIIGVLPF---VRHGPAAKFEKMSASFI--YSVVFVLLACVYGVANNRIHI---VR 158
DB 12 QMGGIFPIYGVTRNDPKRFKMFSLRVILNLTVVVTALLQAYEYGLKAIKINAKVVS 71
QY 159 SLSGPFEEAVIAYLVNLIPIIMIIPIIMYERAKIAKLFNDDEVLVYQISGHSLEPK 218
DB 72 SLIFPIDACLINVLFL-NL-----ATKRSVAMKMDVDDTFNRPYHMQSWS 118
QY 219 LROK-----AVYIAIYLVILSVSVYTHVMSDL-----NIN 251
DB 119 LRKRLGVVSPFLVLAIVEHLISIVSNV--HNQWVEIKYCMWTBPNVFGHYSLRPFANIY 176

```

```

QY 252 QVVPYCIID-----NLTAMIGAMW-----FLICEAMSTIAHLLAERFQKA---LKHIGPA 298
DB 177 LNEFPYNSLSAVFETTVSSALTMWNYODIFIMISIG-----LATRFOQIINNTLKITLSDG 231
QY 299 AMV--ADYRVLMRL-----LSKLTPTDGNALCYTFVFMSTYLFPIITLSYGLMSOLS 349
DB 232 VLIRGDF--WIRVNTVYAVCELDVDYRAISWMLISCATNLYICQILHVSCKLA 288
QY 350 EGFQIXDIGLITLALMNIGLLFYICBAHYASVNVRT-----NFQKCLMVEBLN 398
DB 289 N--TVED-----AYYGFSLGFLI-----VRTVIVFLSAHHIDCAKKPLDIIMK 330
QY 399 WMSDAQTEINMFLRATENMNSTINGGFEDVNRITLFXGLITTVTVYLVLLQFOISIP 458
DB 331 IPNVGMCVELERFSTQKSEKVALSGMGFSLTRQLLFSMAGTIVYELVNLKFD----- 385
QY 459 DKGDSGANNI 469
DB 386 --QESSESKNI 394

```

Search completed: May 23, 2005, 19:25:11  
Job time : 177 secs



**This Page Blank (uspto)**